

7/2/19  
STIC-Biotech/ChemLib

189582

From: Chan, Christina  
Sent: Wednesday, May 10, 2006 5:46 PM  
To: Graser, Jennifer; STIC-Biotech/ChemLib  
Subject: RE: rush search

Please ~~Just~~ Thanks Chris

Chris Chan  
TC 1600 New Hire Training Coordinator and SPE 1644  
(571)-272-0841  
Remsen, 3E89

-----Original Message-----

From: Graser, Jennifer  
Sent: Wednesday, May 10, 2006 5:46 PM  
To: Chan, Christina  
Subject: rush search  
Importance: High

Hi Christina,

Will you please authorize the following rush search a special case which is due?

Thanks,  
Jennifer

-----  
STIC:

Please search nucleotides 11357 to 12736 of SEQ ID NO: 35 from application 10/672787 in pending and commercial databases.

Thanks,  
Jennifer Graser  
Primary Examiner  
Art Unit 1645  
REM 3B09  
2-0858

3C18

NA 96109

mg  
\*\*\*\*\*  
Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*  
Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2006, 08:45:36 ; Search time 8124 Seconds  
(without alignments)  
10862.560 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380  
Sequence: 1 taagcgctgacacgtttaa.....acgacgcacatcttttga 1380

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 11581394

Minimum DB seq length: 22

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_env:\*  
2: gb\_pat:\*  
3: gb\_ph:\*  
4: gb\_pl:\*  
5: gb\_pr:\*  
6: gb\_ro:\*  
7: gb\_sts:\*  
8: gb\_sv:\*  
9: gb\_un:\*  
10: gb\_vi:\*  
11: gb\_ov:\*  
12: gb\_hgt:\*  
13: gb\_in:\*  
14: gb\_om:\*  
15: gb\_da:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 1380   | 100.0       | 96109  | 2     | AR408756    |
| 2          | 1380   | 100.0       | 96109  | 2     | AX067460    |
| 3          | 1370.4 | 99.3        | 1410   | 2     | AR450009    |
| 4          | 266.8  | 19.3        | 110000 | 15    | CP000082_25 |
| 5          | 242    | 17.5        | 110000 | 15    | CR543861_02 |
| 6          | 229.6  | 16.6        | 110000 | 15    | CP000082_24 |
| 7          | 226.2  | 16.4        | 1416   | 2     | AR318732    |
| 8          | 160.4  | 11.6        | 1314   | 2     | AX932204    |
| 9          | 160.4  | 11.6        | 11498  | 2     | AX191762    |
| 10         | 160.4  | 11.6        | 110000 | 2     | BD426631_11 |
| 11         | 160.4  | 11.6        | 110000 | 2     | BD426631_12 |
| 12         | 160.4  | 11.6        | 110000 | 2     | AR274513_11 |
| 13         | 160.4  | 11.6        | 110000 | 2     | AR274513_12 |
| 14         | 160.4  | 11.6        | 110000 | 2     | AR632719_11 |
| 15         | 160.4  | 11.6        | 110000 | 2     | AR632719_12 |
| 16         | 160.4  | 11.6        | 110000 | 15    | L42023_11   |
| 17         | 160.4  | 11.6        | 110000 | 15    | L42023_12   |
| 18         | 157.6  | 11.4        | 110000 | 15    | CR628336_30 |

|    |       |      |        |    |             |                     |
|----|-------|------|--------|----|-------------|---------------------|
| 19 | 157.6 | 11.4 | 110000 | 15 | CR628337_28 | Continuation (29 o  |
| 20 | 154.8 | 11.2 | 110000 | 15 | AE016828_01 | Continuation (2 of  |
| 21 | 154.4 | 11.2 | 1311   | 2  | CS222566    | Sequence            |
| 22 | 154.4 | 11.2 | 14324  | 2  | CQ872956    | Sequence            |
| 23 | 154.4 | 11.2 | 110000 | 15 | CP000057_12 | Continuation (13 o  |
| 24 | 154.4 | 11.2 | 349980 | 2  | CQ873072    | Sequence            |
| 25 | 154.2 | 11.2 | 110000 | 15 | CP000155_59 | Continuation (60 o  |
| 26 | 154.2 | 11.2 | 110000 | 15 | CP000155_60 | Continuation (61 o  |
| 27 | 152.8 | 11.1 | 110000 | 2  | AE017354_29 | Continuation (30 o  |
| 28 | 152.4 | 11.0 | 1314   | 2  | AX932206    | Sequence            |
| 29 | 146.2 | 10.6 | 110000 | 15 | BA000031_04 | Continuation (5 of  |
| 30 | 142.2 | 10.3 | 12232  | 2  | AE006048    | Pasteurel           |
| 31 | 139.6 | 10.1 | 1317   | 2  | AX189058    | Sequence            |
| 32 | 139.6 | 10.1 | 2608   | 15 | ECMUROY     | Escherichia         |
| 33 | 139.6 | 10.1 | 28277  | 2  | AX191720    | Sequence            |
| 34 | 139.6 | 10.1 | 28277  | 15 | EC2MIN      | Escherichia         |
| 35 | 139.6 | 10.1 | 110000 | 15 | U00096_00   | E. coli 2 m         |
| 36 | 139.6 | 10.1 | 110000 | 15 | AE014075_00 | U00096 Escherichia  |
| 37 | 139.6 | 10.1 | 110000 | 15 | AE014075_01 | Continuation (2 of  |
| 38 | 139.6 | 10.1 | 110000 | 15 | AP009048    | Escherich           |
| 39 | 137.4 | 10.0 | 110000 | 15 | AE017340_04 | Continuation (15 of |
| 40 | 136.4 | 9.9  | 110000 | 15 | CP000034_01 | Continuation (12 of |
| 41 | 134.8 | 9.8  | 110000 | 15 | AE005174_00 | AE005174 Escherich  |
| 42 | 134.8 | 9.8  | 110000 | 15 | AE005174_01 | Continuation (2 of  |
| 43 | 134.8 | 9.8  | 110000 | 15 | BA000007_00 | BA000007 Escherich  |
| 44 | 134.8 | 9.8  | 110000 | 15 | BA000007_01 | Continuation (2 of  |
| 45 | 133.2 | 9.7  | 1629   | 15 | ECMURD      | X17609 Escherichia  |

# ALIGNMENTS

|            |  |                                     |     |        |                 |
|------------|--|-------------------------------------|-----|--------|-----------------|
| RESULT 1   | AR408756   | 96109 bp                            | DNA | linear | PAT 18-DEC-2003 |
| LOCUS      | AR408756   | Sequence 35 from patent US 6632636. |     |        |                 |
| DEFINITION | AR408756   |                                     |     |        |                 |
| ACCESSION  | AR408756   |                                     |     |        |                 |
| VERSION    | AR408756.1   | GI:40159151                         |     |        |                 |
| KEYWORDS   |  |                                     |     |        |                 |
| SOURCE     | Unknown.   |                                     |     |        |                 |
| ORGANISM   | Unclassified.  |                                     |     |        |                 |
| REFERENCE  | 1 (bases 1 to 96109)   |                                     |     |        |                 |
| AUTHORS    | Lagace, R.E., Patterson, C. and Berg, K.L.                                 |                                     |     |        |                 |
| TITLE      | Nucleic acids encoding 3-ketoacyl-ACP reductase from Moraxella catarrhalis |                                     |     |        |                 |
| JOURNAL    | Patent: US 6632636-A 35 14-OCT-2003;                                       |                                     |     |        |                 |
| FEATURES   | Elitza Pharmaceuticals Inc.; San Diego, CA                                 |                                     |     |        |                 |
| source     | location/Qualifiers  |                                     |     |        |                 |
|            | 1..96109   |                                     |     |        |                 |
|            | /organism="unknown"  |                                     |     |        |                 |
|            | /mol_type="genomic DNA"  |                                     |     |        |                 |

6632636

|        |  |   |
|--------|--|---|
| ORIGIN | Query Match  | 100.0%; Score 1380; DB 2; Length 96109; |
|        | Best Local Similarity 100.0%; Pred. No. 0;                             |   |
|        | Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;          |   |
| QY     | 1 TAAAGCGTGAACCATGTTAACTTATCCTCGCTCATTAATTTTAAATGATC 60                |   |
| DB     | 11357 TAAAGCGTGAACCATGTTAACTTATCCTCGCTCATTAATTTTAAATGATC 11416         |   |
| QY     | 61 CATTGCTGACCAAGCGTGAACCAAGCGTGAACCAAGCGTGAACCAAGCGTGAAC 120          |   |
| DB     | 11417 CATTGCTGACCAAGCGTGAACCAAGCGTGAACCAAGCGTGAACCAAGCGTGAAC 11476     |   |
| QY     | 121 TACTGCCGATCAAGTGTACCACTGATATATTTAGATCCAAATGTCATACCTTT 180          |   |
| DB     | 11477 TACTGCCGATCAAGTGTACCACTGATATATTTAGATCCAAATGTCATACCTTT 11536      |   |
| QY     | 181 GAGTAATCTTGTCTCAATTAATGCGGCGATTTGACCGATTAATATACACTATCCACCA 240     |   |
| DB     | 11537 GAGTAATCTTGTCTCAATTAATGCGGCGATTTGACCGATTAATATACACTATCCACCA 11596 |   |

241 CTGCATTAACATATGCTGACAGCTCACTAAAATCTTGAACCTTTACCAAGCCCAACCAAAAT 300  
11597 CTGCTAATACATATGCTGACAGCTCACTAAAATCTTGAACCTTTACCAAGCCCAACCAAAAT 11656  
301 CAATGCCAGTGTATGATATGCCATTAACCGTGCAGGAGCCCATCAATTCCTGCAATGTTGA 360  
11657 CAATGCCAGTGTATGATATGCCATTAACCGTGCAGGAGCCCATCAATTCCTGCAATGTTGA 11716  
361 GCCGATGTTGTAACCTTTGAATCAATTAATATGACCTTGTATCAATATCAATCAATATA 420  
11717 GCCGATGTTGTAACCTTTGAATCAATTAATATGACCTTGTATCAATATCAATCAATATA 11776  
421 TTCAACGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGGTATCAAGATGCGATTCATGCGG 480  
11777 TTCAACGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGGTATCAAGATGCGATTCATGCGG 11836  
481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGATGATGCTTGGCC 540  
11837 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGATGATGCTTGGCC 11896  
541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
11897 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11956  
601 ATCAACATCATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
11957 ATCAACATCATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12016  
661 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
12017 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12076  
721 GGCAATGCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
12077 GGCAATGCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12136  
781 CAATAATTCACACATGCGATCGTGCATCAAGATGATGATGATGATGATGATGATGATGATGAT 840  
12137 CAATAATTCACACATGCGATCGTGCATCAAGATGATGATGATGATGATGATGATGATGATGAT 12196  
841 TGCCCTCTGAGCTGCGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
12197 TGCCCTCTGAGCTGCGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 12256  
901 CACCAATTCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
12257 CACCAATTCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12316  
961 GACACCAACATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
12317 GACACCAACATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12376  
1021 TTTGGCATTTGAGAGCTGTGATGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
12377 TTTGGCATTTGAGAGCTGTGATGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 12436  
1081 TGCTTTAGTGTATCAATGAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
12437 TGCTTTAGTGTATCAATGAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12496  
1141 GGGGGGACATATCTGCGCTTTCTTGGGTTAATGCGGGGCTGATGATGATGATGATGATGATGATGAT 1200  
12497 GGGGGGACATATCTGCGCTTTCTTGGGTTAATGCGGGGCTGATGATGATGATGATGATGATGATGAT 12556  
1201 TGAGCTTAAAGTTCATCTGTCAAGCTGCCAAATTAATGATGATGATGATGATGATGATGATGATGAT 1260  
12557 TGAGCTTAAAGTTCATCTGTCAAGCTGCCAAATTAATGATGATGATGATGATGATGATGATGATGAT 12616  
1261 ATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
12617 ATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12676

1321 CAAAAAATTCACAGCTGACAGCTCAGAGCTGCCCAACCGACGACGATATTTTGA 1380  
12677 CAAAAAATTCACAGCTGACAGCTCAGAGCTGCCCAACCGACGACGATATTTTGA 12736

RESULT 2  
AX067460 96109 bp DNA linear PAT 24-JAN-2001  
LOCUS Sequence 35 from Patent WO0078968.  
DEFINITION AX067460  
ACCESSION AX067460  
VERSION AX067460.1 GI:12545080  
KEYWORDS  
SOURCE Moraxella catarrhalis  
ORGANISM Moraxella catarrhalis  
Moraxella catarrhalis  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Moraxellaceae; Moraxella.

REFERENCE  
AUTHORS Lagace, R.E., Patterson, C. and Berg, K.L.  
TITLE Nucleotide sequences of moraxella catarrhalis genome  
JOURNAL Patent: WO 0078968-A 35 28-DEC-2000;  
Incyte Genomics, Inc. (US)

FEATURES  
source Location/Qualifiers  
1..96109  
/organism="Moraxella catarrhalis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:480"

ORIGIN  
Query Match 100.0%; Score 1380; DB 2; Length 96109;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TAAAGCGTGAACATGTTTAACTTAACCTTACCTGCTCATTAATTTTAAATGATC 60  
11357 TAAAGCGTGAACATGTTTAACTTATCACTGCTCATTAATTTTAAATGATC 11416  
61 CATGCTGACACAGCAGTGAACAAAGCACAGCCCTTACCGGCTTGTGCTGCGCAA 120  
11417 CATGCTGACACAGCAGTGAACAAAGCACAGCCCTTACCGGCTTGTGCTGCGCAA 11476  
121 TACGCCCCGATCAAGCTTACCAACCTGATATTTTGAAGCCAAATGTCATCACTTT 180  
11477 TACGCCCCGATCAAGCTTACCAACCTGATATTTTGAAGCCAAATGTCATCACTTT 11536  
181 GAGTAAATCTTGTCAATTTACGGGCGATCTTACCGATTAATTAATCACTATCCACA 240  
11537 GAGTAAATCTTGTCAATTTACGGGCGATCTTACCGATTAATTAATCACTATCCACA 11596  
241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCCAACCAAT 300  
11597 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCCAACCAAT 11656  
301 CAATGCCAGTGTATGATATGCCATTAACCGTGCAGGAGCCCATCAATTCCTGCAATGTTGA 360  
11657 CAATGCCAGTGTATGATATGCCATTAACCGTGCAGGAGCCCATCAATTCCTGCAATGTTGA 11716  
361 GCCGATGTTGTAACCTTTGAATCAATTAATATGCTTGTATCAATATCAATATA 420  
11717 GCCGATGTTGTAACCTTTGAATCAATTAATATGCTTGTATCAATATCAATATA 11776  
421 TTCAACGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGGTATCAAGATGCGATTCATGCGG 480  
11777 TTCAACGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGGTATCAAGATGCGATTCATGCGG 11836  
481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGATGATGATGATGATGATGATGAT 540  
11837 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGATGATGATGATGATGATGATGAT 11896  
541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
11897 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11956  
601 ATCAACATCATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660







QY 1241 TGAACGCTGTGGCAGTTTATCTGCCAAGTAGATTGGATTACATGATGCTTA 1300  
 Db 10057 TCAATAGTAGCAGCAGGCTTTATCTGCTAAATGCGGAGTTCTCTGATTTGTAGTAGCGCA 10116  
 QY 1301 ACTTTGTGACCGTGTATGACCAAAAAATTGACAGCTGACGTCGAGAGCTGCCAAACCG 1360  
 Db 10117 ACTTGTAGCGCTGCTGACGCAATATATGTGACCGATATATCCGAGCTGCCCAAAACCG 10176  
 QY 1361 ACCAGCGCATAT 1372  
 Db 10177 ACCAGCATTTGT 10188

# RESULT 5 CR543861.02/c WPCOMMENT

Sequence split into 36 fragments LOCUS CR543861 Accession CR543861

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| CR543861_00   | 1       | 110000  |
| CR543861_01   | 100001  | 210000  |
| CR543861_02   | 200001  | 310000  |
| CR543861_03   | 300001  | 410000  |
| CR543861_04   | 400001  | 510000  |
| CR543861_05   | 500001  | 610000  |
| CR543861_06   | 600001  | 710000  |
| CR543861_07   | 700001  | 810000  |
| CR543861_08   | 800001  | 910000  |
| CR543861_09   | 900001  | 1010000 |
| CR543861_10   | 1000001 | 1110000 |
| CR543861_11   | 1100001 | 1210000 |
| CR543861_12   | 1200001 | 1310000 |
| CR543861_13   | 1300001 | 1410000 |
| CR543861_14   | 1400001 | 1510000 |
| CR543861_15   | 1500001 | 1610000 |
| CR543861_16   | 1600001 | 1710000 |
| CR543861_17   | 1700001 | 1810000 |
| CR543861_18   | 1800001 | 1910000 |
| CR543861_19   | 1900001 | 2010000 |
| CR543861_20   | 2000001 | 2110000 |
| CR543861_21   | 2100001 | 2210000 |
| CR543861_22   | 2200001 | 2310000 |
| CR543861_23   | 2300001 | 2410000 |
| CR543861_24   | 2400001 | 2510000 |
| CR543861_25   | 2500001 | 2610000 |
| CR543861_26   | 2600001 | 2710000 |
| CR543861_27   | 2700001 | 2810000 |
| CR543861_28   | 2800001 | 2910000 |
| CR543861_29   | 2900001 | 3010000 |
| CR543861_30   | 3000001 | 3110000 |
| CR543861_31   | 3100001 | 3210000 |
| CR543861_32   | 3200001 | 3310000 |
| CR543861_33   | 3300001 | 3410000 |
| CR543861_34   | 3400001 | 3510000 |
| CR543861_35   | 3500001 | 3598621 |

Continuation (3 of 36) of CR543861 from base 200001 (CR543861 Acinetobacter sp. ADP1 cont)

Query Match 17.5%; Score 242; DB 15; Length 110000;  
 Best Local Similarity 52.9%; Pred. No. 3.3e-67;  
 Matches 726; Conservative 0; Mismatches 580; Indels 66; Gaps 7;

QY 5 GCGTGAACCATGTTAACTTAACCTTACCGCTGCTAATTAATTTTAAATGATCAG 64  
 Db 71434 GCAATCACACAGCAAAATTAATGATGCTCCGATGTGATTAACCTTTAAACATATCAAAA 71375  
 QY 65 CTAGACAAGCAGAGTAGCAAAAGACAGCCCTAGCGGTCTTGTGACTGTGCTGCAATACT 124  
 Db 71374 CTTGACAGAGCAGGTGACAAATACAGAGCTTTCAAGGTGT-----GCATACGTGTGT 71321  
 QY 125 GCGGATCAAGGTATACCAACCTGATATATTTTGAATGCCAAATGTGCATACCTTTGAGT 184  
 Db 71320 GCGTGTGTACCGCATCTTTTAAAGTGATGATGACCAACGAGTGTGTTCTGT---C 71264

QY 185 AATCTTGCTCAATTAATCGGGCATCTTGACGATTAATAATACACTATACCACTGC 244  
 Db 71263 AAGCGTGTCAATCTGCTGAGATCTTACCGATACGAAATGCGACCTTCAATATTTTC 71204  
 QY 245 ATACATATGCTGACAGCTCACTAAATCTTGACCTTTAACAAAGCCCAACCAATC--- 301  
 Db 71203 TCAACAGACTGGGCGCAGAGCTAAATATCTGCTTACCTGTGCGCGCCCAAAATCACA 71144  
 QY 302 ---AATGCGAGATGATATGCCATTAACCGTCCCAAGGCCCAATGCTGCATGGTT 358  
 Db 71143 ATGACTTTGCTGTGTGGGTTCAATTGACAGCAACAAACATCAATTCGAGACAGCTT 71084  
 QY 359 GAGCCGATGTTTGTACTCTTTTGAATCATTAATAATGCTGCTATCATATATCATATA 418  
 Db 71083 GCGCCCAATGTGCTCTTTCGATATATATATACGATACATGACTGTGTTTAC 71024  
 QY 419 TATTCACAGGATGCGGTAGCCCTTTGAAAGTCTTAAGGATACAAAGATGCAATC 478  
 Db 71023 TATTCGACAGATGCTCAAGCCCTTTAACTGCTTCAAGGTTTCAAGCATATGACACA 70964  
 QY 479 GGAAGATTGCGAGCTGTGCTAAGGAGTGCAGATTAAGGCAATGATGCTATGCTT 538  
 Db 70963 GGTAAACCTATGCTTACCCCAATGCGACAGACCAAGCATGTCATGCAATATATGTG 70904  
 QY 539 CCTTGATTTTAAATGATGATGTTGTTAAAGTTTGTGTTTGTATGATATAATGCAATA 598  
 Db 70903 CCTGAAATATACAGCTCAAGACTTTAATCAGGCGGATAGGCGCATGTCCAGCCACATG 70844  
 QY 599 CCATACCATCATCAAAATTAATAAAATCTGCCGTTTGTGCTAAGTAAGCTATTTGTT 658  
 Db 70843 GTTCATCATCTTTCAGTAAACGCCATATGCTT-----AATGTGCG 70801  
 QY 659 GTGATATGCTGTGATTTGGGCGAAGCTTGTGACGTGTGATGACAGATGCTTGGTAA 718  
 Db 70800 GTGATGATGATCAAACTTTGATGTGTATCTTGTGACACAGAGGCGGTGTAAC 70741  
 QY 719 TTGCGATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAATATGCAATTTTGT 778  
 Db 70740 T-----GTATACAGGTGTTAAACCACTTTTACACCTGAAATATGATTTAGCC 70685  
 QY 779 GCCAAATTAATCAACATGCGATGTCGATGCAATGATGATGATGATGATGATGATGAT 838  
 Db 70684 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 70625  
 QY 839 GTTGCCCTTGAAGCTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 898  
 Db 70624 ACTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 70565  
 QY 899 ACCACCAATCAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958  
 Db 70564 ACCGACATC---AGCTGCTCTTTTAAACCAATCTAATCTGAGAGCCCTAATTAACCT 70508  
 QY 959 CCGACACCAACATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018  
 Db 70507 CTAATGCAACGAGTTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 70448  
 QY 1019 CTTTGGCATTTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078  
 Db 70447 CTTTGGCATTTGAAACCGTAAATGCAAAATGCAAACTGAGTGTGACGCGCT----- 70394  
 QY 1079 CGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138  
 Db 70393 -----AAGATTGAATATCTCAATGATGATGATGATGATGATGATGATGATGATGAT 70355  
 QY 1139 TTGCGGCAACAATCTGCGCTTCTTGATTAATGCGGGGCTGATGATGATGATGATGATGAT 1198  
 Db 70354 ATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 70295  
 QY 1199 GCTGACCTTAAGATTCATGCTGACAGCTGCAAGATGCAATGCAATGCAATGCAATGCAAT 1258  
 Db 70294 GCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 70235  
 QY 1259 TTATGTCCAAAGTAGATTGGGATTAACATGATGATGATGATGATGATGATGATGATGAT 1318

Db 70234 TGAATGCGCCAGGTGAGTCTCTCGAGATCTGTGACAGCAAGCGGTAGCTTGTTC 70175  
QY 1319 ACCAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAACGACGAGGAT 1370  
Db 70174 TGTAAAAATTAACGCTGAACCCCTGAGATTCGAAGACCCGACGACTT 70123

RESULT 6  
CP000082.24  
WPCOMMENT

Sequence split into 27 fragments LOCUS CP000082 Accession CP000082

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| CP000082.00   | 1       | 11000   |
| CP000082.01   | 100001  | 21000   |
| CP000082.02   | 200001  | 31000   |
| CP000082.03   | 300001  | 41000   |
| CP000082.04   | 400001  | 51000   |
| CP000082.05   | 500001  | 61000   |
| CP000082.06   | 600001  | 71000   |
| CP000082.07   | 700001  | 81000   |
| CP000082.08   | 800001  | 91000   |
| CP000082.09   | 900001  | 101000  |
| CP000082.10   | 1000001 | 111000  |
| CP000082.11   | 1100001 | 121000  |
| CP000082.12   | 1200001 | 131000  |
| CP000082.13   | 1300001 | 141000  |
| CP000082.14   | 1400001 | 151000  |
| CP000082.15   | 1500001 | 161000  |
| CP000082.16   | 1600001 | 171000  |
| CP000082.17   | 1700001 | 181000  |
| CP000082.18   | 1800001 | 191000  |
| CP000082.19   | 1900001 | 201000  |
| CP000082.20   | 2000001 | 211000  |
| CP000082.21   | 2100001 | 221000  |
| CP000082.22   | 2200001 | 231000  |
| CP000082.23   | 2300001 | 241000  |
| CP000082.24   | 2400001 | 251000  |
| CP000082.25   | 2500001 | 261000  |
| CP000082.26   | 2600001 | 2650701 |

Continuation (25 of 27) of CP000082 from base 2400001 (CP000082 Psychrobacter arcticus 2)

Query Match 16.4%; Score 229.6; DB 15; Length 110000;  
Best Local Similarity 54.3%; Pred. No. 4.1e-63;  
Matches 581; Conservative 0; Mismatches 444; Indels 45; Gaps 4;

QY 8 TGAACCATGTTAACAACCTTATCACTCGCTCATTAATTTTAATTGATCCATCTA 67  
Db 108845 TTAACGAGTGGCTTAAGCCCTGCCCCGCGCAAAACCATTAATGATCGAACCTT 108904  
QY 68 GCACAAGCAGGTGACAAAGCAGCCCTAGCGGTGCTTGACTGGCTGCCAATACCTGCC 127  
Db 108905 GCACAAGCAGGTGACAAATATACCGCTGTACTTGGCAGCAAACTACTTGGCTTCTGC 108964  
QY 128 CGA-----TCAAGTATCAACCTGATATATTTAGATGCCAAATGT 169  
Db 108965 TCAATCATCGAAAAAGCATTTTCTAAGTCTGACATGAGCAGGCTAACCTCATGCTG 109024  
QY 170 GCATCACTTGTGTAATCTGTGCTAATATGCGGCATCTTGACCGATTAATATACA 229  
Db 109025 AGACCCGCTCGCTAATCTGTTCATCTGCTGACTGTCTTGCCAATPAAGAGACT 109084  
QY 230 CTATCCACCCACATGACATATGCTGACAGCTCACTAAATCTTACCTTACCAAGC 289  
Db 109085 TGACTCAATATGATGATTAAGGAGGAGTTCACGAAGCTGCTAAGCTTTTACCTTGA 109144  
QY 290 CCAACCCAAATCAATGCCA-----GTGATGATATGCCATPAACCGTGCACAGCCATCA 343  
Db 109145 CCGCCCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 109204  
QY 344 ATTAGCTGCAATGATGATGAGCGGATGTTGACCTTTTAATCAATTAATTAATTAATTAAT 403  
Db 109205 ACAGCCGCAATGATGATGAGCGGATGTTGATGATGATGATGATGATGATGATGATGATG 109264

QY 404 TCAATATCATCATATATATTCACAGCAGTGGGATGCCCTTTGAAGCTTAAGGTATCA 463  
Db 109265 GCATATGTCGCGACATATTTGGCAGCGATGCTCAAGACCTGCAAAACGCTGCAAGGTGCTA 109324  
QY 464 AGCATGGCATCCATGGGAAGATTGGCACTGTGCTTAAAGCAAGTCAATTAAGCATTTG 523  
Db 109325 AGCATCTCTCAAGCGGCAACCGCAAGCTCCCAATGCTCAATCTGCTTGGGATTA 109384  
QY 524 AGTAGTTATGCTGCTTGTGATTTTAAATGATGATGATGATGATGATGATGATGATGATG 583  
Db 109385 AGTAGTTATGACAGCTTTTATCTTAATTTATGTTATGCTGATATGAGAGCTCTGATCCA 109444  
QY 584 TGAATATATGCTATACCATATCACCATCATCAATTAATTAATTAATTAATTAATTAATTA 643  
Db 109445 CGTGCCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 109485  
QY 644 AGTAAGCTATTTGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 703  
Db 109486 ATAGTCCCTTGTGATGAGCATGAATACCAGTGTCAAGTCCGCGCAAACTATCTGACAC 109545  
QY 704 CATGCTGCTAATATTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763  
Db 109546 CAATGCTGCTGCTA--CGATCTTCAAGATTAATCACTACCGATTTGGCAGCTTGGAG 109603  
QY 764 ATACGAATTTTGTGCAAAATATATCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823  
Db 109604 ATACGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109663  
QY 824 GAGAGATTTAAATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
Db 109664 GACATATGAGAGCTGTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109723  
QY 884 CTGACAGCTCAAGCAGCAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 943  
Db 109724 CTGATATATTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109783  
QY 944 GTGCCAATATTAACCGCGCAGCAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1003  
Db 109784 AGCCCAATATTAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109843  
QY 1004 AATGCTGAACAGTCTTTGGCATTTGAGCTTGTGATGCTGATGCTGATGCTGATGCTGATG 1053  
Db 109844 AGCGTGTACAGCTGCTTTTGTGATTTGAGCCAGTATTCGCAACAAATTTG 109893

## RESULT 7

ARJ18732/c ARJ18732 1416 bp DNA linear PAT 17-AUG-2003  
LOCUS  
DEFINITION Sequence 1282 from patent US 6562958.  
ACCESSION ARJ18732  
VERSION ARJ18732.1 GI:33699835  
KEYWORDS

## SOURCE

ORGANISM Unknown.  
Unclassified.REFERENCE 1 (bases 1 to 1416)  
Breton,G. and Bush,D.TITLE Nucleic acid and amino acid sequences relating to Acinetobacter  
baumannii for diagnostics and therapeuticsJOURNAL Patent: US 6562958-A 1282 13-MAY-2003;  
Genome Therapeutics Corporation; Waltham, MAFEATURES  
Location/Qualifiers  
1..1416  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 16.4%; Score 226.2; DB 2; Length 1416;  
Best Local Similarity 52.1%; Pred. No. 2.1e-62;  
Matches 718; Conservative 0; Mismatches 593; Indels 66; Gaps 7;

QY 2 AAAGCGTGAACCATGTTAACAACCTTATCACTCGCTCATTAATTTTAATTAATGATCC 61

```

Db      1409 AAGGAATTGACGAGGCAAACTGCTGACCAAGGTCATTATACCTTTAAACATATCA 1350
Qy      62  ATCTGACACAGAGAGTGACAAAAGACAGCCCTAGCGGTCTTGACGTGCGCAAT 121
Db      1349 AACTTGCACTGCTGCTGAT-----AGCAATACCACTCTTCAGCTTGCTTGCAC 1298
Qy      122  ACGCCGATGCAAGTGTACCACTGATATATTTTGTAGTCCAAAGTGTGATCACTTTG 181
Db      1297 GTTGACACAGCTGACAGCTTTCTTTAAGCGTGTGTCATTAATTTAGTTCGCTT- 1239
Qy      182  AGTAAATCTTGCTCAATTAATCGGGGACCTTGACCCGATTAATATACACTATCCACCAC 241
Db      1238 TGAATGGCTTGTTCATGACAGCGGATCTTCAACATCAATACACAACTTTGGCATAT 1179
Qy      242  TGCATACATATGTCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAAAATC 301
Db      1178 TTTTCATAGAAAGACGTAAAGAAATACTTGTCTTGCCCTGTCCACCTAAATA 1119
Qy      302  AATGCCA-----GTGATGATATGCCATTAACCGTGTCCAGCCCATCAATTTGTCGAATG 355
Db      1118 AGTGACACCTTACCTTTTTCACCTTCAATGACAGACCTAAGCCATCAATTTGTCGAATG 1059
Qy      356  GTTGAGCCGATGTTTGACCTTTTGATTCATTAATAATATGCTTGCTATCAATATCATCA 415
Db      1058 GTGACACCAATATTAGTACCTTTAGATCATTAATTAAGGACATGATGACAGGTTTAA 999
Qy      416  ATATATTCACAGGATGCGGTAGCCCTTTGAAAGCTTTAAGGATATACATGACATGCC 475
Db      998  ACCTTACACAGCGGTGCTCTAGTCTTTAAATTTTAAAGTTTCAACATATGACTCC 939
Qy      476  ATGGGAAGATTGACAGCTGTGCTTAAGGCAAGTGACAGATTAAGGATTTAGATGATATGC 535
Db      938  ATAGTAACCAATGCTTCCCAATGCTAACAAGTAAAGATTAAGCTATCAATGTGC 879
Qy      536  TTGCTTTGATTTTAAATGATGATTTGTAAAGTTTGTTTTGTGATTAATATGCC 595
Db      878  ATACCTTGATATATTAATCTGAGCTTTTAATTAACCGCTTAACCAAGAGCAAGCA 819
Qy      596  ATACCATACCATCATCAATTAATAATAAATTCGCGCTTGTTGGCTTAAGTAAAGCTATTT 655
Db      818  AGCGTACCGTGGCATCTCTTAAACGCCCATATGTAT-----TAAATCGGTGCGT 767
Qy      656  GTTGTATATATGATGATTTGGGGCAAGCTGTGATGTGTCAGATGCTTGCTTGCT 715
Db      766  TTAAACCAAGCTTTGATGATGTGTGATCTGCAACAAATGACGGCTTA----- 716
Qy      716  AAATTTGGCATGCTTTGATGATGATGCGGTTTGCAATTTTCAAAATATACCAATTTT 775
Db      715  ----ACGATCATCTCGGTTAAATACAACTTTTACGCGCTTGGAATAATCATGTGTTT 660
Qy      776  TGTGCCAATTAATTCACATGTCATGTCGCAATCAAGATGATCAAGGAGAGATTTAAA 835
Db      659  GCTTGATGATACCCAGCATATTTCCATGACGGCTTAATGATGCTTCACTCATATTTAGA 600
Qy      836  ATTGTTCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAATTTGACAGTCA 895
Db      599  ACCACGTCACTCAGGCTTTAAGTGAAGGTGTTCTAATGAAAGCTTGATTAATCTCA 540
Qy      896  AGCAGCAACCAATTCATATTTGGATGCTTAATTAATCAAGTCAGAGCGCTGCCAATTTA 955
Db      539  AGGACCAATGATCTGTG-----TTGATCTTTAAGTAATCTTAACGTGTGACCAAGTTG 483
Qy      956  CCGCCGACCAACAAATCATGCTGTCATGTTTGCATCTCGCTACTTAATGTGCTAACA 1015
Db      482  CCGCTTACTGCACTTCTTACCTGCACTCTTACGCTTAAGCCATTAAGTAAAGTGTACA 423
Qy      1016  GTGCTTTGGCATTTGAGCTGTGATGCGATGATGTTGTTGTTGCTTGAGTTAGA 1075
Db      422  GTACTTTTTCATTTGAACTGTAATGCAAAATGCGCAATGCAATGAGTGC----- 373
Qy      1076  TCTGTGCTTTGAGTGTATGATCAATGAAATGCAATGCTGATCAAGGAATGCCCTTGA 1135

```

```

Db      372  -----ACGCGCAATTAATGATATCCGCCACACAGAAATACCTTTA 330
Qy      1136  GCTTTGGCGGACCAATATCTCGCGCTTCTTGCTTATGCGCGGGCTGATGATCTGA 1195
Db      329  GCATATGCTGCTGAAATTTCCGTAATGTGTGGCAAGGCTTGCTTAATAATATCTCT 270
Qy      1196  TCGGCTGACCTTTAAGATTTCATCTGTCCAAGCTGCCAAATGACAAATGACGCTGTGGC 1255
Db      269  TCTGCTTGAATTAATTAATTTTGTATCAAGCTGACCAAACTGCTTTTAACACAGCGCGA 210
Qy      1256  AGTTATCTGCCAAGGTAGATTTGGATTAATCAATGATATGCTAACTTTGACCGTGA 1315
Db      209  ATCTGATCTGTGCTGGAGAGGTGTGGCGGGAATCGTTACAGCACTTGTTAGCTTGT 150
Qy      1316  TTACCAAAAATTCACAGCTGACATCCAGAGCTGCCAAACCGACGACGGATAT 1372
Db      149  TCATGCAAGAAATTTACAGCAAAACCTGATATTTCCAAAGCTGCTCAACACTTTT 93

RESULT 8
AX932204/c 1314 bp DNA linear PAT 22-DEC-2003
LOCUS      AX932204
DEFINITION Sequence 157 from Patent WO03087353.
ACCESSION AX932204
VERSION    AX932204.1 GI:40312625
KEYWORDS
SOURCE
ORGANISM   Haemophilus influenzae
            Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE
  1. Edwards, A., Dharamsi, A., Vedula, M., Domagala, M., Houston, S.,
    Awrey, D., Beatrice, B., Mansoury, K., Ouyang, H., Veljee, F.,
    Richards, D., Neherly, K., Viray, C., Banadja, K., Pinder, B.,
    Alam, M.Z., Tal, M., Canadien, V., and Kanagarajah, D.
    Patent: WO 03087353-A 157 23-OCT-2003;
    Affinium Pharmaceuticals, Inc. (CA)
FEATURES
  source
    1..1314
    /organism="Haemophilus influenzae"
    /mol_type="unassigned DNA"
    /db_xref="taxon:727"
ORIGIN
Query Match      11.6%; Score 160.4; DB 2; Length 1314;
Best Local Similarity 49.6%; Pred.No.1,2e-40;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;
Qy      1  TAAAGCGTGAACCATGTTAAACAACTTATCACTGCTCATTAATTTTAATTGATC 60
Db      1308 TAAACATTGAGCTTAATGCTTAATTTCTTGCGCGCTTTTCAAAAGAACAACTGATC 1249
Qy      61  CATGCTAGCACAAGAGGTGACAAAAGACAGCCCTAGCGGTCTTTGACGTGCGCAA 120
Db      1248 GAGACTTGACACAAGCGGACGACATATATCA-----TATCTCGCTTGCAA 1201
Qy      121  TACTGCCGATCAAGGTATCAACCTGATATATTTTATAGTCCAAATGTGATCACTTT 180
Db      1200 TGTGGCGGTAAAAATTTATATGCTTGTTCATTTGATGCAACAA-----ATPACTTTG 1147
Qy      181  GAGTAATCTTGCTCAATTAATCGGGCATCTTGACCGATTAATTAATACACTATACCCA 240
Db      1146 CGATGAAAATTTTGGACAGCGCCACATCTGACAAACAAATTAATATATATGTGTG 1087
Qy      241  CTGCATPAACATATGCTGACAGCTCACTAAATCTTAACTTACCAAGCCACCAAAAT 300
Db      1086 TTGATTAATTAATTCAGCTAATCTTAAATAATGAGCCCTTTTCGCTCTCGCTGCAA 1027
Qy      301  CAATGCAAGTATGATATGCAATAACCGTCCAAAGCCATCAATTTGCTGCAATGTGTA 360
Db      1026 CAATGCAATTAACCTCAATATA-----AAAGCCAGCCAAATGACCAACTGACT 976
Qy      361  GCCGATGTTTGTACTTTGATCATTAATAATATGCTTGCTATCAATATCATCAATATA 420

```

|    |      |   |      |
|----|------|---|------|
| Db | 975  | CCCCACATTTTGCTGTAGAGTCATTATATCCAAAGATGCGATTAGCTTATGCACTAA     | 916  |
| Qy | 421  | TTCAAGCGATGCGGTAACCCCTTGAAGTCTTAAAGGTATCAAGCATGCGATCCATGGG    | 480  |
| Db | 915  | TTGAAAACGATGATCTAACCCCTTGAATATGAGAAAGTCGGATGGAATTTGAATCTAAATT | 856  |
| Qy | 481  | AAGATGGCAGCTGTGCTTAAGGCAAGGACAGATTAAGGCATTGAGTAGTATGCTTGCC    | 540  |
| Db | 855  | AATACCTATACCTTGTGCGCAATGCTGTGTGCGCAAAATGTTCAATATATATGCGACG    | 796  |
| Qy | 541  | TTTGATTTTAAATGATGGAATGGTAAAGTTGTGTTTTTGAATGATTAATGCCATACC     | 600  |
| Db | 795  | AACCAATGATCCTTCTTCAAGAGTAATACATTCATCTTTTACATTAATATATATGCTT    | 736  |
| Qy | 601  | ATCACATTCATCAATAATAATAAAATCGC--CGTTGGTGCTAGTAGCTATTTGTTG      | 659  |
| Db | 735  | GCCATTTTCACTTTTGAAGCCAAATATCCGACATATTTTCCGAAAGAAACGATGTGTT    | 676  |
| Qy | 660  | TGATATATGTTGATTTGGGGCAACGCTGTCTAGTGTGTCAAGCATGCTTGGCTAAAT     | 719  |
| Db | 675  | CGCTGATTTTGGTTTTCCCAAAAGTAGGC-----                            | 644  |
| Qy | 720  | TGGCATGCTTGTGGTATGATGAGGGGTTTGCAATTTTCAAAATACGCAATTTTGTG      | 779  |
| Db | 643  | ---TATCTTCAATGTTCACACACCTACTTAGCTATATGATTAATATGTTATTTGCTT     | 587  |
| Qy | 780  | CCAATATATCACCGATGCGATCGTGCGCATMAAGATGATCAGCGAGAGATTTAAATATG   | 839  |
| Db | 586  | GCGCATATCTTCTTAATCCATATAGCATTCATATGATCTTCACTACGTTCAAGACAG     | 527  |
| Qy | 840  | TTGCCCTTGAAGTGCACAAATTTGATATGCTCAAGTGAATCTTGACAGCTTCAAGCA     | 899  |
| Db | 526  | TGCGTGCAGCGACTTTTAAAGCTATAAAGTGTCTCAAGCTGAATCTAATAAGCTTAACTA  | 467  |
| Qy | 900  | CCACCAATATCATTTTGTGATTCGTATATTTAAAGTGAAGGGGTGCGCATATTAACCG    | 959  |
| Db | 466  | CATTAAGTTCACATCTTTCATTC--AACATATGACAAAGCGGGAATCCCAATATTTCCG   | 410  |
| Qy | 960  | CGACACCAACATCATGCTGTCATGTTTGGCATCTCGCTACTATATGTCGTAAACAGTGC   | 1019 |
| Db | 409  | CCATACCAACTTTACACACACAGCTTTCCGCATTTCTAATACTPAATAGTTACGGTAC    | 350  |
| Qy | 1020 | TTTTGGCATTTGAGCTGTGATGGCGATGATTTGGTGTGTTGTGCTTGAAGTAACTTC     | 1079 |
| Db | 349  | TTTAAACATTTGAACCTGTAAATCCCAAAATTTGGTTGTGCTGCGGCGCA-----       | 298  |
| Qy | 1080 | GTGCTTTGAGTATCAATGAACAATTAACATGCTGATACAGAAATGCTTGAAGCTT       | 1139 |
| Db | 297  | -----GAATATTCATATATGCGATTAATCTTCCACTCCCGCTTAA                 | 257  |
| Qy | 1140 | TGGGGGCGACAATCTCGGCGTTCTTGCGGTTAATCCGGGGCTGATGATCTGATGCG      | 1199 |
| Db | 256  | GTGGGGTTTGAATTTCTGTGTTTTTTTACCGCAAGCTTGGGCTATAAACAATCATATGCG  | 197  |
| Qy | 1200 | CTGAGCTTAAGTCACTGTCTCAAGCTGCCAAATGAACATGAAGCGCTGCGGCAGTT      | 1259 |
| Db | 196  | TTTCAAGTACATCTTCTGATTTTAACATCCAGTATGAAGGGATATTTTGAAGAACTT     | 137  |
| Qy | 1260 | TATCTGCAGAGTAGATTTGGGATTTACCATGATGATGCTAACTTTGTACCGTGATTTGA   | 1319 |
| Db | 136  | TATCAATACCAAGTAGATTTTTTGGAGTATCAATACACAGAAATATTAGCTGTGGGATTA  | 77   |
| Qy | 1320 | CCAAAAATTTCAAGCTACAGTCCAGAGCTGCCAAACGACGACGCGATATTTTTT        | 1377 |
| Db | 76   | AGAGTAAATCCACAAAGAAAGACCTGTGTTTGGCAAGCCCAATGATTAATATTTTTT     | 19   |

|            |            |                                    |                        |
|------------|------------|------------------------------------|------------------------|
| RESULT 9   |            |                                    |                        |
| LOCUS      | AX191762/c |                                    |                        |
| DEFINITION | AX191762   | 11498 bp                           | DNA                    |
| ACCESSION  | AX191762   | Sequence 44 from Patent WO0149775. | linear PAT 15-AUG-2001 |

|           |   |             |
|-----------|---|-------------|
| VERSION   | AX191762.1  | GI:15209931 |
| KEYWORDS  |   |             |
| SOURCE    |   |             |
| ORGANISM  | Haemophilus influenzae<br>Haemophilus influenzae<br>Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;<br>Pasteurellaceae; Haemophilus. |             |
| REFERENCE | 1   |             |
| AUTHORS   | Iversen, P.L.   |             |
| TITLE     | Antisense antibacterial cell division composition and method  |             |
| JOURNAL   | Patent: WO 0149775-A 44 12-JUL-2001;  |             |
| FEATURES  | Avi Biopharma, Inc. (US)<br>Location/Qualifiers   |             |
| source    | 1..1148<br>/organism="Haemophilus influenzae"<br>/mol_type="unassigned DNA"<br>/db_xref="taxon:727"   |             |
| ORIGIN    |   |             |

| Query Match | Similarity | 11.6%   | Score 160.4       | DB 2       | Length 11498 |
|-------------|------------|---|-------------------|------------|--------------|
| Best Local  | Similarity | 49.6%   | Pred. No. 1.9e-40 |            |              |
| Matches     | 683        | Conservative  | 0                 | Mismatches | 606          |
|             |            |   |                   | Indels     | 89           |
|             |            |   |                   | Gaps       | 77           |
| Qy          | 1          | TAAAGCGTGAACCATGTTAACAACTTATACCTGCTCATATTAATTTTAAATGATC       | 60                |            |              |
| Db          | 9122       | TAAACATGAGCTAAATGCTGAATTTCTTCGCCGCGCTTTTCAAAAGAGCAAACTGATC    | 9063              |            |              |
| Qy          | 61         | CATGCTAGCAAAACAGGTGACAAAGACAGCCCTAGCGGTGCTTGACTGGCTGCCAA      | 120               |            |              |
| Db          | 9062       | GAGACTTGCAACAGCAGCGGACAAATATATCA-----TATCTCGCTTTGCA           | 9015              |            |              |
| Qy          | 121        | TATCGCCCGATCAAGTGTACCAACCTGTATATTTTATAGATCCAAATGTGATCACTTT    | 180               |            |              |
| Db          | 9014       | TGTTGGGGGTAAATTTCTATCGCTGTTGTCATTTGATGCAACA-----ATATCTTGG     | 8966              |            |              |
| Qy          | 181        | GAGTAAATCTTGCTCATATTATGCGGGCATCTTGACCGATTAAATGACATATCCACCA    | 240               |            |              |
| Db          | 8960       | CGATGAAATTTTGTGACAGCAGCGCACCTCTCGACCAAAACAAATAAATATGTGTG      | 8903              |            |              |
| Qy          | 241        | CTGCACTACATATGTGTGACAGCTCACTAAATCTTGAACCTTACCAAGCCACCCAAAT    | 300               |            |              |
| Db          | 8900       | TTGATTAATTAATTCAGCTAATTTCTGAAAATTCAGCCCTTTTCCGTCTCCGCTTAGCA   | 8844              |            |              |
| Qy          | 301        | CAATGCCAGTATGATATGACCATTAACCGTGCACAGCCCATCAATTTGTCGAATGTGA    | 360               |            |              |
| Db          | 8840       | CAATATGCAATTTACCTCATATTA-----AAGCCACAGCATGACAGCACTGTACT       | 8799              |            |              |
| Qy          | 361        | GCCGATGTTTGTACTCTTTTGAATCATTTAAATATATGCTTGTATCAATATCATCAATTA  | 420               |            |              |
| Db          | 8789       | CCCCACATTTTGTGTTTGAATCATTAATCCACGAATGTGCATTAGCTTATGTGACTTA    | 8733              |            |              |
| Qy          | 421        | TTCAAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGATCCATGG    | 480               |            |              |
| Db          | 8729       | TTGAAAAGCATGATCTAACCTTTTGAAATGACGAAGTGGCGGTGCAATTTGAATCTAAT   | 8676              |            |              |
| Qy          | 481        | AAGATGGCAGCTGTGCTTAAAGCAGTGCAGATTAAGCATTTAGTATGTATGCTTCC      | 540               |            |              |
| Db          | 8669       | AATACCTTATGCTTGTGCGCAATGCTGTGCTGCACAAATGTTCATATTAATATGCGCAC   | 8610              |            |              |
| Qy          | 541        | TTTGATTTTAAATGATGATTTGTTGAAGTAAGTTTGTTTTGTATGATATTAATGCATAC   | 600               |            |              |
| Db          | 8609       | AACCAATGTAGCTTCTTCAAGGTAATAATCATTCATCTTTTACCATTAATATATGCTT    | 8556              |            |              |
| Qy          | 601        | ATCACCATCATCAATAATAATAAATATGCG-CGTTGGTGGCTAAGTAACTATTTGTG     | 659               |            |              |
| Db          | 8549       | GCCATTTTCAATTTTATAGCCAAATATCCGCACTATTTTCCGAAAGAAACGATATGTT    | 8496              |            |              |
| Qy          | 660        | TGATTAATGTGATTTGGGCAACGCTTGTGATGTGTGATCAGCATTTGCTTGCTAAAT     | 719               |            |              |
| Db          | 8489       | CGCTTGATTTTGTGTTTCCCAAAAGTCAGC-----                           | 8458              |            |              |
| Qy          | 720        | TGCAATGCTCTTGATGATGATGCGGCTTGGCAATTTTCAAAATAGCAATTTTGG        | 779               |            |              |
| Db          | 8457       | ---TATCTTCATTTGTTCAACACACTTACTTTAGCATTAATGAATTAATGCGTATTTTGCT | 8401              |            |              |



|    |      |   |      |
|----|------|---|------|
| QY | 780  | CCAAATTAATCCACCATGCCATCGCTGCCGATCAAGTATGATCAGCGGAGAGATTTTAAATTG | 839  |
| Db | 8400 | GGCGCAATTAATCTTCTMAATATCCATATATGAGCATTCATATGATCTTCAAGTCAAGG     | 8341 |
| QY | 840  | TTGCCCTCTGAGCTGCCAAATTTTGATATGCTCAAGCTGAAACTTACAGCTCAAGCA       | 899  |
| Db | 8340 | TGCGTCCGCGAGCTTTTAAAGCTATATAGTGTCTCAAGCTGAAACTTAAAGCTCTAATA     | 8281 |
| QY | 900  | CCACCAATTCATATTTTGGATTCGTTAATTAATCAAGTCAGCGCTGCCAATATTAACGC     | 959  |
| Db | 8280 | CATAAAGTTCAAAATCTTCATTC---AACAATGACAAGCGGGAATCCCAATATTTCCGC     | 8224 |
| QY | 960  | CGACACCAACATCATCGCTGCATCTTTGCCATCTCGGCTAATATGTGTAACAGTC         | 1019 |
| Db | 8223 | CCATACCAACTTTCACACACGAGGCTTTGCGCATTTCTAATACTMAATAGTTTCGATAC     | 8164 |
| QY | 1020 | TTTTGGCATTTGAGCTGTGATGCGGATGATTTGGTGTGTTGTCCTGAGTTAGATTC        | 1079 |
| Db | 8163 | TTTTTACATTTGAACTGTAAATCCCCAAATTTGGCTTTGTGCTGCGGCGCA-----        | 8112 |
| QY | 1080 | GTGCTTTGAGTGTATCAATGAAACAATTGAAATCGCTGATCAAGAAATGCTTGAGCTT      | 1139 |
| Db | 8111 | -----GAATATTTCAATATGCGCATTAATCTTCCACTCCGCTTTAA                  | 8071 |
| QY | 1140 | TGCGGCGCACAATACTCGGCGTTCTTGCGGTTAATCCGGGCTGATGATGATCTGATCGG     | 1199 |
| Db | 8070 | GTGCGGTTTGAATTTCTGTGTTTTTAAACCGAAGCCCTGGGCTAATAAACAATCATATCGC   | 8011 |
| QY | 1200 | CTGAGCTTAAAGTTCACATGTCACAGCTGCCAAATGCAATGAAACGCTGTGGGCACTT      | 1259 |
| Db | 8010 | TTTCAAGTACCATTCCTGATTTTAACTACCAAGTATGAAAGGGATATTTTGAAGGAAGTT    | 7951 |
| QY | 1260 | TATCTGCCAAGTATGAGATTTGGGATTAACATCGATGATCTTAATCTTTGACACGATATGA   | 1319 |
| Db | 7950 | TATCAATATCCAGTAGAATTTTTTGCAGATATCAATACACGAATATATAGCCGTGGGATA    | 7891 |
| QY | 1320 | CCAAAAATTCACAGCTGACAGTCCAGCTGCCAAACGACGACGCGCATATTTTTT          | 1377 |
| Db | 7890 | AGAAATATATCAACAAGAAAGACCTGTTTTGGCAAGCCCATATATGATATATTTTT        | 7833 |

| Sequence split into 19 fragments |         |         |                                   |
|----------------------------------|---------|---------|-----------------------------------|
| Fragment Name                    | Begin   | End     | LOCUS BD426631 Accession BD426631 |
| BD426631_00                      | 1       | 110000  |                                   |
| BD426631_01                      | 100001  | 210000  |                                   |
| BD426631_02                      | 200001  | 310000  |                                   |
| BD426631_03                      | 300001  | 410000  |                                   |
| BD426631_04                      | 400001  | 510000  |                                   |
| BD426631_05                      | 500001  | 610000  |                                   |
| BD426631_06                      | 600001  | 710000  |                                   |
| BD426631_07                      | 700001  | 810000  |                                   |
| BD426631_08                      | 800001  | 910000  |                                   |
| BD426631_09                      | 900001  | 1010000 |                                   |
| BD426631_10                      | 1000001 | 1110000 |                                   |
| BD426631_11                      | 1100001 | 1210000 |                                   |
| BD426631_12                      | 1200001 | 1310000 |                                   |
| BD426631_13                      | 1300001 | 1410000 |                                   |
| BD426631_14                      | 1400001 | 1510000 |                                   |
| BD426631_15                      | 1500001 | 1610000 |                                   |
| BD426631_16                      | 1600001 | 1710000 |                                   |
| BD426631_17                      | 1700001 | 1810000 |                                   |
| BD426631_18                      | 1800001 | 1830121 |                                   |

Continuation (12 of 19) of BD426631 from base 1100001 (BD426631 The Nucleotide Sequence)

Query Match 11.6%; Score 160.4; DB 2; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 3e-40;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
1 TAAAGCGTAACCATGTTAACAACTTATCATCTCGCTCATTAATATTTTAAATGATC 60

|    |        |  |        |
|----|--------|--|--------|
| Db | 104439 | TAAACATTAGCTAAATGCGTAAATCTTTCGCGCGCTTTCAAAAAGAAACAACTGATC      | 104380 |
| Oy | 61     | CATGCTACACACAGAGGTGACAAAGAACAAGCCCTAGCGGTGCTTGACTGGCTCCAA      | 1120   |
| Db | 104379 | GAGATTTCACAGACAGCGGCAATATACCA-----TATCTCCGCTTGCAA              | 104332 |
| Oy | 121    | TACTGCCGATCAAGTGTAACCAACTGTATATTTTGAATGCCAAATGTGCATCACCTTT     | 180    |
| Db | 104331 | TGTTGGGGGTAAAAATTTCTATGCTGTGTTCCATGTATGCAACA-----ATACCTTG      | 104278 |
| Oy | 181    | GAGTAAATCTGCTCAATTTATTCGGGGCATCTTGACCGATTAATATACACTATCCACCA    | 240    |
| Db | 104277 | CGATGAATAATTTTGGACAGCGCACCATCTGACCAAAACAATACAAATATATGTGTG      | 104218 |
| Oy | 241    | CTGCATTAACATATGCTGACAGCTCACTAAAATCTTGAACCTTATACAAAGCCCAAAAT    | 300    |
| Db | 104217 | TTGATTAATTAATTACGCTAATCTTGAAAAATCAGCCCTTTCCGTCCTCGCTAGCAA      | 104158 |
| Oy | 301    | CAATGCCAGTGAATGATGCCATTAACCGTGCAAGCCCATCAATGTCTGCAATGCTTGA     | 360    |
| Db | 104157 | CAATAGCAATTTACCTCAATATTA-----AAGCCAGCAATGACAGCAACTGTACT        | 104107 |
| Oy | 361    | GCCGATGTTTGAACCTTTTGAATCATTTAAATATGCTTGCATATCAATATCATATTA      | 420    |
| Db | 104106 | CCCCCAATTTGTGTTTGAAGTCAATTAATCCAGAAATGCCATTAAGCTTGATGACCTAA    | 104047 |
| Oy | 421    | TTACAGCGATGCGGTAGCCCTTTGAAAAGTCTTAAGGGATCAAGCATGGACCTCATGGG    | 480    |
| Db | 104046 | TTGAAAACGATGATTAACCTTTTAAATAGCAAGATGGCGATACGAATGAATCTAAAT      | 103988 |
| Oy | 481    | AAGATTGCAGCTGTGCTTAAGGCAAGTGACAGATTAAGGCAATGAGGTATATGCTTCC     | 540    |
| Db | 103986 | AATACCTATAGCTGTGCAATGCTGTGCTGCCAAAATGTTCAATATATATATGCGAGC      | 103927 |
| Oy | 541    | TTTGATTTTAAATGATGAGATTTGTTAAAGTTGTTTTTGAATGATATTAATGCCATACC    | 600    |
| Db | 103926 | AACCATGTAGCTTCTTCAACAAGTAAATCATCTTATCTTTTACCATTAAATATTTGCTT    | 103867 |
| Oy | 601    | ATACACATCATCAATTAATATAAAATCTGC--CGTTGGGCGTAAGTAAATGATTTGTTG    | 660    |
| Db | 103866 | GCCATTTTCAAGTTTATAGCCAAATATCCGACTATTTTCCGAAAGAAACGGTATGTTT     | 103807 |
| Oy | 660    | TGATATGATGTGATTTTGGGCAACGCTGTCTGTCAGTGTGCTCAAGCATTCCTGGCTAAAT  | 719    |
| Db | 103806 | CGCTGATTTTCGTTTTCCCAAAGTCAAGC-----                             | 103775 |
| Oy | 720    | TGGCATGCTCTTGTGATATGATGGCGGTTTGGCAATTTTCAAAAATACCAATTTTGTG     | 779    |
| Db | 103774 | ---TATCTTCAATGTTTCAACACACCTACTTGTAGCATTTATGATTAATGCGTAATTTGCTT | 103718 |
| Oy | 780    | CCAAATATTCACCAATGTCATGCTGTCCCATCAAGATGATCAGGGAGAAATTTAAATTG    | 839    |
| Db | 103717 | GCGGTAATCTTCTTAATTCATATAGCATATGATCTTACGTCAGCTCAACAG            | 103658 |
| Oy | 840    | TTGCCCTTGAAGCTCCAAATTTGATATGCTCAAGCTGAAACCTTGCAGCTCAAGCA       | 899    |
| Db | 103657 | TGCGTCGCGCAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAACTGAAGCTCTAGTA     | 103598 |
| Oy | 900    | CCACCAATTCATATTTTGGATCGTTAATTAATTTCAAGTGACGGCGTGCCAAATTTACCG   | 959    |
| Db | 103597 | CATAAAGTTCAAAATCTTCATTC---AACAAATGACAAGCGGAATCCCAATATTTCCCG    | 103541 |
| Oy | 960    | CGACACCAACATCATGCTGTCATGTTTGGCATCTGCGCTATCTAATGTCTGAACAGTGC    | 1019   |
| Db | 103540 | CCATATCCAACTTTCACACACAGAGCTTTGCGCATTTTCATTAACATAAAGTATACGCTAC  | 103481 |
| Oy | 1020   | TTTTCGCAATTTGAGCTGTGATGCGGATGATGATGTTTGTGCTTATAGTATGATCTC      | 1079   |
| Db | 103480 | TTTTCACATTTGAACCTGTAAATCCCACAATTTGGCTTGTGCTGCGCGGCA-----       | 103429 |
| Oy | 1080   | GTGCTTGAAGTATCAATGAACAATTGAACATGCTGATATCACAGAAATGCTTGAGCTT     | 1139   |

103428 -----GAATTAATTCAATATCGCGATTACTCCACTCCGCTTAA 103388  
QY 1140 TGGCGCGCAATACACTCGGCTTCTGGTTAATGCGGGGCTGATGATGATCGG 1139  
DB 103387 GTGCGGTTGAATTTCTGGGTTTAAACCGCAAGCCCTGGCTTAATACATATATCGC 103328  
QY 1200 CTGAGCTTAAGAGTTCACTGTCCAGCTGCGCAAAATGACAATGACGCTGTGGGAGTT 1259  
DB 103327 TTTCAAGTAACATTCCTGATTTAACTACAGATGAAAGGATATTTTGAAGAATT 103268  
QY 1260 TATCTGCCAAGTATGATTTGGGATTTACATCGATGATGCTAATTTGTGACCGTATGA 1319  
DB 103267 TATCAATACCAAGATGATTTTTCGATGATCAACAGAAATTAAGCTGTGGGATA 103208  
QY 1320 CCATAAATTCACAGCTGACAGTCCAGAGCTGCCAAACGACGAGCATATTTT 1377  
DB 103207 AGAGATATTCACACAGAAAGACTGTTTGGCCAGCCGATGATGATATTTT 103150

RESULT 11  
BD426631\_12/c  
WPCOMMENT

Sequence split into 19 fragments LOCUS BD426631 Accession BD426631

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| BD426631_00   | 1       | 110000  |
| BD426631_01   | 100001  | 210000  |
| BD426631_02   | 200001  | 310000  |
| BD426631_03   | 300001  | 410000  |
| BD426631_04   | 400001  | 510000  |
| BD426631_05   | 500001  | 610000  |
| BD426631_06   | 600001  | 710000  |
| BD426631_07   | 700001  | 810000  |
| BD426631_08   | 800001  | 910000  |
| BD426631_09   | 900001  | 1010000 |
| BD426631_10   | 1000001 | 1110000 |
| BD426631_11   | 1100001 | 1210000 |
| BD426631_12   | 1200001 | 1310000 |
| BD426631_13   | 1300001 | 1410000 |
| BD426631_14   | 1400001 | 1510000 |
| BD426631_15   | 1500001 | 1610000 |
| BD426631_16   | 1600001 | 1710000 |
| BD426631_17   | 1700001 | 1810000 |
| BD426631_18   | 1800001 | 1830121 |

Continuation (13 of 19) of BD426631 from base 1200001 (BD426631 The Nucleotide Sequence)

Query Match 11.6%; Score 160.4; DB 2; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 38-40;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGGTGAACCATGTAAACAATTATCAGCTGCTCATTTAATTTTAAATTGATC 60  
DB 4439 TAAACATTGAGTAAATGCTAAATCTTCGCGCTTTTCAAAAGAACAACTGATC 4380  
QY 61 CATGCTAGCAACAGCGGTGACAAAGACAGCCCTAGCGGTCTTGAGTGGCTCCAA 120  
DB 4379 GAACTTGAACAAGCGGACAAATATACCA-----TATCTCGCTTTGCAA 4332  
QY 121 TACTGCCGATCAAGTATCAACACTGATATTTTATAGATGCCAAATGTGATCCTTT 180  
DB 4331 TGTGGGCGTAAATTTCTATCGCTTTCATTTGATGACAA-----ATATCTTTG 4278  
QY 181 GAGTAAATCTTGTCAATTATGCGGGCATCTGACCGATTAATATACACTATACCCA 240  
DB 4277 CGATGAAATTTTGAAGCGACGACCATCTGACCAAAACAATTAATATGTGTG 4218  
QY 241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 300  
DB 4217 TTGATTAATTAATTCAGCTAATTTCTGAAATATCAGCCCTTTTCCTCGCTTGCAA 4158  
QY 301 CAATGCCAGTATGATATGCCATAAACCGTGCAAGCCCACTCAATTTGCTCAATGGTTGA 360  
DB 4157 CAATATCATTTTACCTCAATAT-----AAGCCCAAGCCCAATGCAACCACTGTACT 4107

QY 361 GCGAGTGTGTACCTTTTGAATCATTTAAATATGCTTGTATCATATCATATATA 420  
DB 4106 CCCACATTTGTGTCTTTAGATCATTAATCCAAAGCAATGCAATGCTTGTATGACTAA 4047  
QY 421 TTCACGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGATATCAAGATGCGATCCATGG 480  
DB 4046 TTGAAACATATGATTCACCTTTTGAATGACGAATGCGGTACGATGAATCTAAAT 3987  
QY 481 AAGATTGGCAGTGTGCTTAAGGCAAGTACAGATGAGCATTTAGAGTTATGCTTGGC 540  
DB 3986 AATACCTATATGCTTGTGCGCAATGCTTGTCTGCAAAAGCTTCAATATATATGGGACC 3927  
QY 541 TTTGATTTTAAATGATGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 600  
DB 3926 AACCAATATGCTTCTTCAAGGTAAATCACTTATCATCTTTTACCATTAATATGCTT 3867  
QY 601 ATACACATATCAATAAATTAATAATATGTC -GTTTGGTGTATAGTATTTGTTG 659  
DB 3866 GCCATTTTCAAGTTTATGACCAATATATGCACTATTTTCGCAAAAGAAACGATGTT 3807  
QY 660 TGAATATGATGATTTGGGCAAGCTTGTCAAGTGTGTCAGCATTTGCTGCTAAT 719  
DB 3806 CGCTTATTTTGTGTTTCCCAAAAGTACGC----- 3775  
QY 720 TGGCATGTCTTGTATGATGATGCGGTTTGGCAATTTTCAAAATACCAATTTTGTG 779  
DB 3774 ---TATCTTATTTGTCAACACACTTATGATGATGATGATGATGATGATGATGAT 3718  
QY 780 CCATAATATCCACCATGCGATGCGGATCAAGATGATGATGATGATGATGATGATGATG 839  
DB 3717 GCGATATATCTTAAATCATATATGCGATCCATATGATGATGATGATGATGATGATGAT 3658  
QY 840 TTGCCCCCTTGAAGTGCATATTTGAGATATGCTCAAGCTGAAACCTTGACAGTCAAGA 899  
DB 3657 TCGTCCCGCAGCTTTTAAGCTATAAGTGTCTCAAGCTGAAACCTGATGATGATGAT 3598  
QY 900 CCAACAAATCATATTTTGAATGCTTAATTAATCAAGTACGCGGCTGCAATTTACCGC 959  
DB 3597 CATTAATGTCACATCTTATTC---ACAAATGACAAACGCGGAATCCCAATATTTCCGC 3541  
QY 960 CCAACAAATCATATGCTGATGCTTGGCAATGCTGCGCTCTATATGCTATGATGATG 1019  
DB 3540 CCAATACAACTTTCACACAGCAGCTTTCGCAATTTCAATTAATTAATGATGATGATG 3481  
QY 1020 TTTTGCATTTGAGCTGTGATGCGATGATGATGATGATGATGATGATGATGATGATG 1079  
DB 3480 TTTTACATTTTGAACCTGATATCCCAATATGCTTGTGCTGCGGCA----- 3429  
QY 1080 GTGCTTGTAGTATCATATGAACAATTTGAACATGCTGATCACAGGAATGCTTGAAGCTT 1139  
DB 3428 -----GAATTAATTAATATATGCGCGATTAATCTTCCACTCCGCTTTAA 3388  
QY 1140 TGGCGCGCAATACACTCGGCTTCTGGTTAATGCGGGGCTGATGATGATGATGATG 1199  
DB 3387 GTGCGGTTGAATTTCTGATGTTTAAACCGCAAGCCCTGAGCTTAATCAATCATATCGC 3328  
QY 1200 CTGAGCTTAAGAGTTCACTGTCCAGCTGCAAAATGACAAATGACAGCCTGTGGGAGTT 1259  
DB 3327 TTTCAAGTAAACATTCCTGATTTAACTACAGTATGAAGAGGATATTTTGAAGAATT 3268  
QY 1260 TATCTGCCAAGTATGATTTGGGATTTACATCGATGATGCTAATTTTGTGACCGTATGA 1319  
DB 3267 TATCAATACCAAGATGATTTTTCGATGATCAATCAAGAAATATTAAGCTGTGGGATA 3208  
QY 1320 CCATAAATTCACAGCTGACGTCAGCTGCGCAAGCCCAACGACGAGCATATTTT 1377  
DB 3207 AGAGATATTCACACAGAAAGACTGTTTGGCAAGCCGATGATGATGATATTTT 3150

RESULT 12  
AR274513\_11/c  
WPCOMMENT  
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513



Fragment Name      Begin      End  
 AR274513\_00      1      110000  
 AR274513\_01      100001      210000  
 AR274513\_02      200001      310000  
 AR274513\_03      300001      410000  
 AR274513\_04      400001      510000  
 AR274513\_05      500001      610000  
 AR274513\_06      600001      710000  
 AR274513\_07      700001      810000  
 AR274513\_08      800001      910000  
 AR274513\_09      900001      1010000  
 AR274513\_10      1000001      1110000  
 AR274513\_11      1100001      1210000  
 AR274513\_12      1200001      1310000  
 AR274513\_13      1300001      1410000  
 AR274513\_14      1400001      1510000  
 AR274513\_15      1500001      1610000  
 AR274513\_16      1600001      1710000  
 AR274513\_17      1700001      1810000  
 AR274513\_18      1800001      1830121

Continuation (12 of 19) of AR274513 from base 1100001 (AR274513 Sequence 1 from patent US 6,111,111)

Query Match      11.6%;      Score 160.4;      DB 2;      Length 110000;  
 Best Local Similarity 49.6%;      Pred. No. 3e-40;  
 Matches 683;      Conservative 0;      Mismatches 606;      Indels 89;      Gaps 7;

1 TAAAGGCTGACCATGTTAAACAATCTATCACTCGCTCTATTAATTTTAAATTGATC 60  
 104439 TAAACATTTGAGCTTAATGCGTAAATCTTCGCGCGTTTAAAGAAACAACGATC 104380  
 61 CATGCTAGCAGACAGAGTGAACAAGACAGCCCTAGCGGCTGCTTGAAGCTGCGCA 120  
 104379 GAACTTTCAGACAGAGGAGGACCAATATACCA-----TATCTCGCTTGCA 104332  
 121 TACTGCCGATCAAGTGTACCACTGATATATTTAGATGCCAAATGTGCATCACCTT 180  
 104331 TGTGGGCGTAAATAATTCATCGCTTGTTCATTTGATGACAA-----ATTACTTTG 104278  
 181 GAGTAAATCTGTCTCAATTTATGCGGGCATTTGACCGATTAATACACATATCCACCA 240  
 104277 CGATGAAATTTTTCAGAGCGCCACCATCTGCACCAAAACAAATTAATATGTGTG 104218  
 241 CTCGCTAACATATGCTGACAGCTCACTAAATCTTGACCTTTCAGAGCCCAAAAT 300  
 104217 TTGATTAATTAATTCAGCTAATCTGAAAAATCAGCCCTTTCCTGCTCGCTAGCA 104158  
 301 CAATGCCAGTATGATATGCCATAAACCGTGCAGAGCCCATCAATTGCTGCAATGTGA 360  
 104157 CAATGCAATTTTACCCCATATTA-----AAGCCAGCAAGTGCAGCACTGTACT 104107  
 361 GCCGATGTTTACCTTTTGAATCAATAATATGCTTGTCTATCAATATCATATATA 420  
 104106 CCCCAATTTGTTGCTTATAGCATTAATCAACGAATCCATTAAGCTTGAATGACTAA 104047  
 421 TTACAGCGATGCGGTAGCCCTTGAAGCTTAAGGGTATCAAGCATGSCATCCAGG 480  
 104046 TTAAAAAGATGATCTAACCTTTGAATGACGAAGTGCAGATTAATTAATTTAATT 103987  
 481 AAGATTGCGAGCTGTGCTTAAGGCAAGTGCAGATTAAGCATTAAGTATGCTTGC 540  
 103986 AATACCTATAGCTTGTGCGCAATGCTGTGCTGCAGAAATGCTATTAATTAATGCGACC 103927  
 541 TTTGATTTTAAATGATGATTTGTTTATGTTTATTTTATGATGATTAATGCCATACC 600  
 103926 AACCAATGATGCTTCTTCAAGGTAAATCACTTCACTTTTAACTTAATTAATTTGCTT 103867  
 601 ATCAACATCACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 659  
 103866 GCAATTTTCACTTTTATGCGCAATTAATTAATTAATTAATTAATTAATTAATTTG 103807  
 660 TGAATATGATGATTTGCGGCAACGCTTGTCAATGATGATGATGATGATGATGAT 719  
 103806 CGCTGATTTTCTGTTTCCCAAAAGTCAACC----- 103775

720 TGGCATGCTTGTGATGATGAGCGGTTGGCAATTTTCAAAATATACCAATTTTGTG 779  
 103774 ---TATCTTCAATTTTCAACACACTTACTTATGATTAATTAATTTTGTCTT 103718  
 780 CCAATTAATCAACATGCAATGCTGCGATCAAGATGATCAGCGAGATTTAAATTTG 839  
 103717 GGGATTAATCTTCAATCAATATAGCATTCATATGATCTTCAATGATCAAG 103658  
 840 TTGCCCCCTGAGCTGCGCAATTTGAGATATGCTCAAGCTGAAATCTTGACAGTCAAGCA 899  
 103657 TCGCTGCGCAGCTTTTAAAGCTATATGATGCTCAAGCTGAAATCTGAAAGCTCAAGTA 103598  
 900 CCAACCAATCAATTTTGTGATGCTTAATTAATTAATTAATTAATTAATTAATTTACCCG 959  
 103597 CATTAATTCATATTTTCAATTCATTC-----AACAAATGACAAAGCGGAAATCCCAATTTTCCCG 103541  
 960 CGACACCAACAAATCAATGCTGCAATGTTTTCGCAATCTGCTTCTAAATGCTPAACAGTGC 1019  
 103540 CCAATCAATTTTCAACACAGAGCTTTCGCAATTTTCAATTAATTAATTAATTAATTTACGCTAC 103481  
 1020 TTTTGGCAATTTGAGCTGCTGATGCGATGATGATGATGATGATGATGATGATGATGAT 1079  
 103480 TTTTACCAATTTGAACTGTAATCCCAATTTGCTTGTGCTGCGCGCA----- 103429  
 1080 GTGCTTGAATGATCAATGAACAAATTTGAACATGCTGATCAAGAGATGCTTGAGCTT 1139  
 103428 -----GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAA 103388  
 1140 TGGCGGCAATTAATCTGCGGCTTCTGCTGTTAAATGCGGGGCTGATGATGATGATGATGAT 1199  
 103387 GTGCGGTTGAATTTCTGAGTTTAAACCGCAAGCCCTGCGGCTTAATTAATTAATTAATTAAT 103328  
 1200 CTGAGCTTAAGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259  
 103327 TTTCAAGTAAACCAATCTGATTTAACTACAGTATGAAGAGGATTTTGAAGAGTT 103268  
 1260 TATCTGCGCAAGGATTTGAGATTTACATGATGATGATGATGATGATGATGATGATGATGAT 1319  
 103267 TATCAATTAATTAATTAATTTTTCGATTAATTAATTAATTAATTAATTAATTAATTAAT 103208  
 1320 CCATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1377  
 103207 AGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103150

RESULT 13  
 AR274513\_12/c  
 WPCOMMENT

Sequence split into 19 fragments      LOCUS AR274513 Accession AR274513

Fragment Name      Begin      End  
 AR274513\_00      1      110000  
 AR274513\_01      100001      210000  
 AR274513\_02      200001      310000  
 AR274513\_03      300001      410000  
 AR274513\_04      400001      510000  
 AR274513\_05      500001      610000  
 AR274513\_06      600001      710000  
 AR274513\_07      700001      810000  
 AR274513\_08      800001      910000  
 AR274513\_09      900001      1010000  
 AR274513\_10      1000001      1110000  
 AR274513\_11      1100001      1210000  
 AR274513\_12      1200001      1310000  
 AR274513\_13      1300001      1410000  
 AR274513\_14      1400001      1510000  
 AR274513\_15      1500001      1610000  
 AR274513\_16      1600001      1710000  
 AR274513\_17      1700001      1810000  
 AR274513\_18      1800001      1830121

Continuation (13 of 19) of AR274513 from base 1200001 (AR274513 Sequence 1 from patent US 6,111,111)

Query Match      11.6%;      Score 160.4;      DB 2;      Length 110000;

Best Local Similarity 49.6%; Pred. No. 3e-40;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGGTGACCATGTTAACTATATACCGTCTCTTAAATTTTAATGATC 60  
DB 4439 TAAACATTTGAGCTAAATGCTAAATCTTCCGCGCTTTTCAAAAGCAAACTGATC 4380  
QY 61 CATGCTAGCAAGACAGGTGACAAAGACAGCCCTTGAACGCTGCGCA 120  
DB 4379 GAACTTTGACAAAGAGGAGCAATATATCA-----TATCTCGCTTGCA 4332  
QY 121 TACTGCCGATCAAGTATCAAACTGATATATTTTATGATGCCAATGTGATCCTTT 180  
DB 4331 TGTGGGCGTAAATAATCTATCTGCTGTTCAATGTATCGAACAA-----ATACTTTG 4278  
QY 181 GAGTAATCTTGTCAATTTATGGGGGACTTGACCGATTAATATACATATACCA 240  
DB 4277 CGATGAAAAATTTTGCAAGACGCGACCATCTGACAAAACAAATTAATATGTGTGG 4218  
QY 241 CTGCTATACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCCAAAAT 300  
DB 4217 TTGATTAATTAATTCAGCTAATTCGAAAAATCAGCCCTTTTCGGTCTCGCTAGCA 4158  
QY 301 CAAATGCCAGTATGATATGCCATPAACCGTCCAAAGCCCATCAATTGCTGCAATGTTGA 360  
DB 4157 CAAATGCAATTTTACCCCTCAATATA-----AAGCCAGCAAGTGCAGCAACTGTACT 4107  
QY 361 GCGGATGTTTGAACCTTTTAATCACTAAATATGCTTCTATCAATATCAATATA 420  
DB 4106 CCCACATTTGTTGCTTAAAGTATTAATCAACGAAATCCATTAAGCTTGAATGCACTAA 4047  
QY 421 TTCAACAGCATGCGGTACCCCTTTGAAGCTTAAAGGATTCACAGATGCGATCATGG 480  
DB 4046 TTGAAAAAGATGATCTTAACCTTTGAATGACGAAGTGGGTGATGATTAATTAAT 3987  
QY 481 AAGATTGGACGCTGTGCTTAAGCAAGTGAATAGCAATGATAGTTATGCTTCC 540  
DB 3986 AATACCTATATGCTGTGCGCAATGCTGTGCTGCCAAATGTTCAATTAATATGCGAC 3927  
QY 541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
DB 3926 AACCATATGATGCTTCTACAGATTAATCACTTCACTTAAATTAATGCTT 3867  
QY 601 ATCACCATCATCAATTAATTAATTAATCTGC-CGTTGTGCTGCTAATGCTATTTGTTG 659  
DB 3866 GCCATTTTCACTTTTAAGCCATATATCCGACATATTTTCGCAAAAGAAACGATATGTT 3807  
QY 660 TGAATATGATGATTTGGGGCAACGCTTGTCAATGTGTCAAGCATGCTTGGCTAAT 719  
DB 3806 CGCTGATTTTCTGTTTCCCAAAAGTCAAGC----- 3775  
QY 720 TGGCATGCTTGTGATGATGATGCGGTTGGCAATTTTCAAAATPACGAATTTTGTG 779  
DB 3774 ---TATCTTATTTGTTCAACACACTATCTTACATTAATATATATGCGTAATTTGCTT 3718  
QY 780 CCAATATATCAACCATGCGCATGCTGCGCATCAAGATGATGCGGAGATTTTAAATTTG 839  
DB 3717 GGGGATATCTTAAATCAATATAGGATCAATATATCTTCAAGTCAAGTCAAGCAG 3658  
QY 840 TTGCCCCCTGAGTCCCAATTTGAGATATGCTCAAGCTGAAAACCTTGACAGCTCAAGCA 899  
DB 3657 TCCCTGCGGAGCTTTTAAGCTATATGTTGTCTCAAGCTGAAAACCTTGAAGCTTATGTA 3598  
QY 900 CCACCAATGCTATTTTGTGATGTTAAATTAATCAAGTGAAGCGGAGCAATTTTACCGC 959  
DB 3597 CATTAAGTTTACATCTTCACTTC--AACAATGACAAAGCGGAAATCCCAATATTTCCG 3541  
QY 960 CGACACCAACATATGCTGATGTTTGTGCTGCTGCTTAAATGTGTAACAGTGC 1019  
DB 3540 CCAATACCAATTTTCAACGAGGAGCTTTCGCAATTTTCAATTAAGTATGATACGTAC 3481  
QY 1020 TTTTGGCATTTGAGCTGTGATGCGCATGATGTGTGTGTTGCTTGAGTTAGATCTC 1079

DB 3480 TTTTACCAATTTGAACTGTATATCCCAATATGCTTTGTGCTGCGCGCA----- 3429  
QY 1080 GTGCTTTGATGTATCATATGAACATTTGAAATGCGGTATCACAGGAATGCTTGAGCTT 1139  
DB 3428 -----GAATTAATTCATATGCGGATTAATCTTCACTCCGCTTTAA 3388  
QY 1140 TGGCGGCGAATATACCTGCGCTTTCTGGGTAAATGCGGGGCTGATGATGATGCGG 1199  
DB 3387 GTGCGGTTGAATTTCTGAGTTTATACGCAAGCCCTGGGCTAATTAACATATATTCG 3328  
QY 1200 CTGAGCTTAAAGATTACATGCTCCAGCTGCCAAATATGACAAATGAAAGCCTGTGGCAGTT 1259  
DB 3327 TTTCAGTAAACCAATTCCTATTTAACTACAGTAAAGAGGATATTTTGAAGAAT 3268  
QY 1260 TATCTGCCAAGTATGAGATTGGGATTTACATGATGATGCTTATCTTGTGACCGTGAATGA 1319  
DB 3267 TATCAATACCAATGAGATTTTTCGATATCAATCAAGAAATTAATGCTTGTGGGAAT 3208  
QY 1320 CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCAAACCGACGACGAGCATATTTT 1377  
DB 3207 AGAGATATTCACAAAGAAAGACTGTTTGTGCAAGCCCGATGATATATATTTT 3150

# RESULT 14 AR632719\_11/c WPCOMMENT

Sequene split into 19 fragments LOCUS AR632719 Accession AR632719

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| AR632719_00   | 1       | 110000  |
| AR632719_01   | 100001  | 210000  |
| AR632719_02   | 200001  | 310000  |
| AR632719_03   | 300001  | 410000  |
| AR632719_04   | 400001  | 510000  |
| AR632719_05   | 500001  | 610000  |
| AR632719_06   | 600001  | 710000  |
| AR632719_07   | 700001  | 810000  |
| AR632719_08   | 800001  | 910000  |
| AR632719_09   | 900001  | 1010000 |
| AR632719_10   | 1000001 | 1110000 |
| AR632719_11   | 1100001 | 1210000 |
| AR632719_12   | 1200001 | 1310000 |
| AR632719_13   | 1300001 | 1410000 |
| AR632719_14   | 1400001 | 1510000 |
| AR632719_15   | 1500001 | 1610000 |
| AR632719_16   | 1600001 | 1710000 |
| AR632719_17   | 1700001 | 1810000 |
| AR632719_18   | 1800001 | 1830121 |

Continuation (12 of 19) of AR632719 From base 1100001 (AR632719 Sequence 1 from patent US

Query Match 11.6%; Score 160.4; DB 2; Length 110000;

Best Local Similarity 49.6%; Pred. No. 3e-40;

Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
QY 1 TAAAGGTGACCATGTTAACTATATACCGTCTCTTAAATTTTAATGATC 60  
DB 104439 TAAACATTTGAGCTAAATGCTAAATCTTCCGCGCTTTTCAAAAGCAAACTGATC 104380  
QY 61 CATGCTAGCAAGACAGGTGACAAAGACAGCCCTTGAACGCTGCGCA 120  
DB 104379 GAACTTTGACAAAGAGGAGCAATATATCA-----TATCTCGCTTGCA 104332  
QY 121 TACTGCCGATCAAGTATCAAACTGATATATTTTATGATGCCAATGTGATCACTTT 180  
DB 104331 TGTGGGCGTAAATAATTTATATGCTTGTTCATATGATGAAACA-----ATACTTTG 104278  
QY 181 GAGTAATCTTGTCTCAATTTATGGGGGACTTGAACGATTAATATACACTATACCA 240  
DB 104277 CGATGAAAAATTTTGCAAGACGCGACCATCTGACAAAACATATATATATGTGTG 104218  
QY 241 CTGCTATACATATGCTGACAGCTCACTAAATCTTGAACCTTATACCAAGCCCAAAAT 300  
DB 104217 TTGATTAATTAATTCAGCTAATTTGAAAAATCAGCCCTTTTCCGCTCGCTAGCA 104158

QY 301 CAATGCCAGTATGATATGCCATTAACCGTGCACAGCCCAATCAATGCTGCAATGTTGA 360  
Db 104157 CAATGCAATTTACCCATTAATATATATATATATATATATATATATATATATATAT 104107  
QY 361 GCCGATGTTGACCTTTTGAATCATTAATATATATATATATATATATATATATATAT 420  
Db 104106 CCCCACTTTGTTGCTTTAGATCATTAATCAAGATGCCATTAATGATGACCTAA 104047  
QY 421 TTCAAGCGATGGGATGAGCTTTGAAAGCTTAAAGGTAATCAAGCATGCGATCATGGG 480  
Db 104046 TTGAAAAGATGATCTAACCTTTGAAATGACGAAGTCCGTACGAATGATCTAAAT 103987  
QY 481 AAGATTGCACTGTGCTTAAGCAAGTGCAGATTAAGCAATGAGTATGCTTGGC 540  
Db 103986 AATACATAGCTGTGCTGCAATGCTGTGCGCAAAATGTTATATATATATATAT 103927  
QY 541 TTGATTTTAAATGATGATGTTGTAAGTTGTTTGTGTTGATGATTAATGCCATACC 600  
Db 103926 AACCAATGATGCTTTCACAAAGTAAATCACTTCACTTTTACATTAATATGCTT 103867  
QY 601 ATGACCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 659  
Db 103866 GCCATTTTCACTTTTGAACCAATTAATCGACATATTTCCGCAAAAGAAACGTAAT 103807  
QY 660 TGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719  
Db 103806 CGCTGATTTTCTGTTTCCCAAAAGTCAAGC----- 103775  
QY 720 TGGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779  
Db 103774 ---TATCTCATTTGTTCAACACACCTTATGATGATGATGATGATGATGATGATGAT 103718  
QY 780 CCAATTAATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839  
Db 103717 GGGCATATCTTCTTAATTCATTAATGATGATGATGATGATGATGATGATGATGAT 103658  
QY 840 TTGCGCTTGAAGTGCCTTAATTTGATGATGATGATGATGATGATGATGATGATGAT 899  
Db 103657 TCGTGGCGGAGCTTTTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 103598  
QY 900 CCAACCAATCAATTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 959  
Db 103597 CATTAATGATCAATCTTCAATTC---AACAATGACAAAGCGGGAATCCAAATATTC 103541  
QY 960 CGACACCAACATATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
Db 103540 CCAATACCACTTTCACACGACGACCTTCCGCAATTTATTAAGTATTAAGTAT 103481  
QY 1020 TTTTGGCATTTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
Db 103480 TTTTACATTTGAACTGTGATATCCCAATTTGCTTGTGCTTGAATGATCTC 103429  
QY 1080 GTGCTTGAATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1139  
Db 103428 -----GAATTAATCAATATGCGGATTAATCTCACTCCCGCTTAA 103388  
QY 1140 TGGGCGGACAAATGCTGCTTCTTGGGTTAATGCGGGGCTGATGATGATGATGAT 1199  
Db 103387 GTGGGCTTGAATTTCTGGGTTTAACTCGCAACCTTGGGCTTAATTAATCAATTA 103328  
QY 1200 CTGAGCTTAAGATTCATGCTCAAGCTGCCTAAATGACATGAACCTGTGTGGCAGT 1259  
Db 103327 TTTCAATTAACATTTCTGATTTTAATTAATTAATTAATTAATTAATTAATTAAT 103268  
QY 1260 TATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319  
Db 103267 TATCAATACAGATGATGATTTTTCGATGATCAATCAACGAATTAATGCTGTGGATA 103208  
QY 1320 CCAAAATTCAGCTGACGATGACAGCTGCCCAACGACGACGATATTTT 1377  
Db 103207 AGAATATCAACCAAGAAAGACCTGTTTGGCAAGCCGATATGATTAATTTT 103150

RESULT 15  
AR632719\_12/c  
MPCOMMENT  
Sequence split into 19 fragments LOCUS AR632719 Accession AR632719  
Fragment Name Begin End  
AR632719\_00 1 110000  
AR632719\_01 100001 210000  
AR632719\_02 200001 310000  
AR632719\_03 300001 410000  
AR632719\_04 400001 510000  
AR632719\_05 500001 610000  
AR632719\_06 600001 710000  
AR632719\_07 700001 810000  
AR632719\_08 800001 910000  
AR632719\_09 900001 1010000  
AR632719\_10 1000001 1110000  
AR632719\_11 1100001 1210000  
AR632719\_12 1200001 1310000  
AR632719\_13 1300001 1410000  
AR632719\_14 1400001 1510000  
AR632719\_15 1500001 1610000  
AR632719\_16 1600001 1710000  
AR632719\_17 1700001 1810000  
AR632719\_18 1800001 1830121  
Continuation (13 of 19) of AR632719 from base 1200001 (AR632719 Sequence 1 from patent UK  
Query Match 11.6%; Score 160.4; DB 2; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 3e-40;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
QY 1 TAAAGCTGAACATGTTAACTTAATCAATCACTTACCTGCTCATTAATTTTAAATGATC 60  
Db 4439 TAAACATGAGCTTAATGCTTAATTTCTTGGCGGCTTTTCAAAAGCAAACTGATC 4380  
QY 61 CATGTAAGCAAGACAGTGAACAAAGACAGCCCTAGCCGCTTGTACTGCTGCCAA 120  
Db 4379 GAGACTTGCAAGACAGGCGCAATATATACCA-----TATCTCCGCTTGCAA 4332  
QY 121 TACTGCCCATCAAGTATACCACTGATATATTTAGATGCCAAATGTCATCACTTT 180  
Db 4331 TGTGGGCGTAAATTTCTATGCTTGTTCATGATGATGACAA-----ATTAATTTG 4278  
QY 181 GAGTAATCTTCTCAATTAATGCGGCGATCTGACGATTAATTAATCACTATCAACCA 240  
Db 4277 CGATGAATTTTTCAGACGCGCACATCTCAACCAAACTTAATTAATGCTG 4218  
QY 241 CTGATTAATGATGCTGACGCTCACTAAATCTTGAACCTTTCAGACCCACCAAAAT 300  
Db 4217 TTGATTAATTAATTCAGCTAATCTGAAATATCAGCCCTTTTCCGCTCCGCTAGCAA 4158  
QY 301 CAATGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Db 4157 CAATGCAATTTTACCTCAATTA-----AAGCCGACCAAGTACCACTGACT 4107  
QY 361 GCGATGTTTGAATCTTTTGAATCAATTAATTAATGCTTCTATCAATATCATATATA 420  
Db 4106 CCCCACTTTGTTGCTTTTAAGTCAATTAATCAACGAATGCTTGAATGATGATTA 4047  
QY 421 TTCAAGCGATGCGGATGCCCTTTGAAAGCTTAAAGGTAATCAAGCATGCAATCCAT 480  
Db 4046 TTGAAAAGATGATCTTAACCTTTGAAATGACCAAGTCCGCTGCAATTAATTAAT 3987  
QY 481 AAGATTGCACTGTGCTTAAGCAAGTGCAGATTAAGCATTAAGTATGATGATGATGAT 540  
Db 3986 AATTAATGCTTGTGCTCAATGCTGTGCTGCAAAATGTTATTAATTAATTAATG 3927  
QY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 3926 AACCAATGATGCTTTCACAAAGTAAATCACTTATCTTTTACATTAATTAATGCTT 3867  
QY 601 ATCAACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 659  
Db 3866 GCATTTTCAATTTTAAAGCAATTAATCCGCACTATTTTCCGCAAAAGAAACGATGTT 3807





CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1398 BP; 408 A; 306 C; 295 G; 389 T; 0 U; 0 Other;

Query Match 100.0%; Score 1380; DB 8; Length 1398;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAAGCGTGAACCATGTTAACAACCTTATCAGCTGCTCATTTAAATTTTAAATTGATC 60  
DB 1383 TAAAGCGTGAACCATGTTAACAACCTTATCAGCTGCTCATTTAAATTTTAAATTGATC 1324  
QY 61 CATGCTAGACCAAGGAGGTGACAAAGACAGCCCTTAGCGGTCTTGACGTGCTGCCAA 120  
DB 1323 CATGCTAGACCAAGGAGGTGACAAAGACAGCCCTTAGCGGTCTTGACGTGCTGCCAA 1264  
QY 121 TACTGCCGATCAAGTGTACCAACTGATATATTTTAAAGTGAATGTCATCACCTT 180  
DB 1263 TACTGCCGATCAAGTGTACCAACTGATATATTTTAAAGTGAATGTCATCACCTT 1204  
QY 181 GAGTAAATCTTGTCTCAATTAATCGGGGCACTTTGACCGATTTAAATACACTATCCACC 240  
DB 1203 GAGTAAATCTTGTCTCAATTAATCGGGGCACTTTGACCGATTTAAATACACTATCCACC 1144  
QY 241 CTGCAATACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 300  
DB 1143 CTGCAATACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 1084  
QY 301 CAATGCCAGTATGATATGTCATTAACCGTGCACAGCCCATCAATTTGCTGCATGTTGA 360  
DB 1083 CAATGCCAGTATGATATGTCATTAACCGTGCACAGCCCATCAATTTGCTGCATGTTGA 1024  
QY 361 GCCGATGTTTGAACCTTTGAATCATTAATAATGCTTGTCTATCAATATCATCAATATA 420  
DB 1023 GCCGATGTTTGAACCTTTGAATCATTAATAATGCTTGTCTATCAATATCATCAATATA 964  
QY 421 TTCACAGCGATGCGGTGAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGTCATCGGG 480  
DB 963 TTCACAGCGATGCGGTGAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGTCATCGGG 904  
QY 481 AAGATTGACAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGGTTATGCTTGC 540  
DB 903 AAGATTGACAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGGTTATGCTTGC 844

QY 541 TTTGATTTTAAATGATGATGTTGTAAGAGTTGTTTTTGTGATGATATATGCCATACC 600  
DB 843 TTTGATTTTAAATGATGATGTTGTAAGAGTTGTTTTTGTGATGATATATGCCATACC 784  
QY 601 ATACCATCATCAATTAATAATAATCTGCGCTTGTGCTGAGTAACTATTGTTGT 660  
DB 783 ATACCATCATCAATTAATAATAATCTGCGCTTGTGCTGAGTAACTATTGTTGT 724  
QY 661 GATATGATGATGTTGGGGCAAGCGTGTGAGTGTGCTCAAGCATTTGCTGGCTAAAT 720  
DB 723 GATATGATGATGTTGGGGCAAGCGTGTGAGTGTGCTCAAGCATTTGCTGGCTAAAT 664  
QY 721 GGCATGCTTGTGTAGATGATGAGCGGTTTGCAATTTTGAATAATACGAATTTTGTG 780  
DB 663 GGCATGCTTGTGTAGATGATGAGCGGTTTGCAATTTTGAATAATACGAATTTTGTG 604  
QY 781 CAATATATCCACCATGCCATCGTCCGATCAAGATATCAAGCGAGAGATTTAAATGT 840  
DB 603 CAATATATCCACCATGCCATCGTCCGATCAAGATATCAAGCGAGAGATTTAAATGT 544  
QY 841 TGCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAATTGACAGCTCAAGAC 900  
DB 543 TGCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAATTGACAGCTCAAGAC 484  
QY 901 CACCAATCCATATTTTGTGATGTTAATTAATTAATCAAGTGAAGCGCTGCCAATATTA 960  
DB 483 CACCAATCCATATTTTGTGATGTTAATTAATTAATCAAGTGAAGCGCTGCCAATATTA 424  
QY 961 GACACCAACATATGATGCTGATGTTTGTGCTATCTGCTCACTAATGTGTAAACATGCT 1020  
DB 423 GACACCAACATATGATGCTGATGTTTGTGCTATCTGCTCACTAATGTGTAAACATGCT 364  
QY 1021 TTTGCGATTTGAGCTGTGATGAGCGATGATGAGTGTGTTGTTGCTTGAAGTATCTG 1080  
DB 363 TTTGCGATTTGAGCTGTGATGAGCGATGATGAGTGTGTTGTTGCTTGAAGTATCTG 304  
QY 1081 TGCTTTAGTGTATCAATGAATGAACATGCTGATCAGAGAAATCCTTGAGCTTT 1140  
DB 303 TGCTTTAGTGTATCAATGAATGAACATGCTGATCAGAGAAATCCTTGAGCTTT 244  
QY 1141 GGGCGGACAAATATCTGGGCTTCTTGCGTTAATGCCGGGGCTGATGATCTGATCGGC 1200  
DB 243 GGGCGGACAAATATCTGGGCTTCTTGCGTTAATGCCGGGGCTGATGATCTGATCGGC 184  
QY 1201 TGAAGTTAAGATTCATGCTGCAAGCTGCCAAATGACAAATGAAGGCCCTGCGGAGTT 1260  
DB 183 TGAAGTTAAGATTCATGCTGCAAGCTGCCAAATGACAAATGAAGGCCCTGCGGAGTT 124  
QY 1261 ATCTGCCAAGTATGATGAGTGAATTAACATGATGATGCTTAATCTTGTGACCGTATGAC 1320  
DB 123 ATCTGCCAAGTATGATGAGTGAATTAACATGATGATGCTTAATCTTGTGACCGTATGAC 64  
QY 1321 CAAAAAATTCACAGCTGACAGTTCAGAGCTGCCAAACCGACGACGCGCATATTTTGTGA 1380  
DB 63 CAAAAAATTCACAGCTGACAGTTCAGAGCTGCCAAACCGACGACGCGCATATTTTGTGA 4  
RESULT 2  
AAF28548  
ID AAF28548 standard; DNA; 96109 BP.  
XX  
XX AAF28548;  
XX  
XX 04-APR-2001 (first entry)  
XX  
XX Genomic fragment #35.  
XX  
XX Genomic library; bacteria; human upper airway; ocular media; sinusitis;  
XX  
XX bronchopulmonary; endocarditis; meningitis; ss.  
XX  
XX Moraxella catarrhalis.  
XX  
XX  
PN WO200078968-A2.

XX 28-DEC-2000.  
 PD 16-JUN-2000; 2000MO-US016649.  
 PF 18-JUN-1999; 99US-0140121P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX Lagace RE, Paterson C, Berg KL;  
 PI WPI; 2001-041427/05.  
 XX Genomic library for identifying diagnostic and therapeutic compositions,  
 PT and for identifying virulence factors, regulatory elements and drug  
 PT targets, comprises Moraxella catarrhalis nucleic acids.  
 XX Claim 1; Page 345-368; 545DP; English.  
 XX The present invention relates to a Moraxella catarrhalis genomic library  
 CC comprising of a combination of 41 nucleic acid molecules (see AAF28514-  
 CC AAF28554). The library has a number of uses described in the  
 CC specification e.g. is useful for identifying diagnostic and therapeutic  
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large  
 CC aerobic, gram-negative diplococcus, normally found among the bacterial  
 CC flora of human upper airways. M. catarrhalis is known to cause acute,  
 CC localised infections such as otitis media, sinusitis and bronchopulmonary  
 CC infection and life-threatening, systemic diseases including endocarditis  
 CC and meningitis  
 XX  
 XX Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;  
 QY Query Match 100.0%; Score 1380; DB 4; Length 96109;  
 Db Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;  
 Matches 1380; Conservative 0; Indels 0; Gaps 0;  
 QY 1 TAAAGCGTGAACATGTTAAACAACTATCATCCTGCTCATTTAATTTTAAATGATC 60  
 Db 11357 TAAAGCGTGAACATGTTAAACAACTATCATCCTGCTCATTTAATTTTAAATGATC 11416  
 QY 61 CATGCTGACACAGAGGTGACAAAGACAGCCCTAGCGGTGCTTGAAGTGGCCAA 120  
 Db 11417 CATGCTGACACAGAGGTGACAAAGACAGCCCTAGCGGTGCTTGAAGTGGCCAA 11476  
 QY 121 TACTGCCCCATCAAGTGTACCAACCTGATATATTTTGAAGCCAAATGTGATCCTTT 180  
 Db 11477 TACTGCCCCATCAAGTGTACCAACCTGATATATTTTGAAGCCAAATGTGATCCTTT 11536  
 QY 181 GAGTAAATCTTGTCAATTAATCGGGGACCTGTGACCGATTAATAACACATATCCACCA 240  
 Db 11537 GAGTAAATCTTGTCAATTAATCGGGGACCTGTGACCGATTAATAACACATATCCACCA 11596  
 QY 241 CTGCAATACATATGTGACAGCTCATTAATCTTTGACCTTTTACCAAGCCACCAAAAT 300  
 Db 11597 CTGCAATACATATGTGACAGCTCATTAATCTTTGACCTTTTACCAAGCCACCAAAAT 11656  
 QY 301 CAATGCCAGTGAATATGCGCATTAACCGTGGCCAGCCCAATCTTCTGCAATGCTTGA 360  
 Db 11657 CAATGCCAGTGAATATGCGCATTAACCGTGGCCAGCCCAATCTTCTGCAATGCTTGA 11716  
 QY 361 GCCGATGTTGTACTTTTGAATCATTAATAATGCTTGTCTATCAATATCATCAATATA 420  
 Db 11717 GCCGATGTTGTACTTTTGAATCATTAATAATGCTTGTCTATCAATATCATCAATATA 11776  
 QY 421 TTGACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGATGCGATCCATGG 480  
 Db 11777 TTGACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGATGCGATCCATGG 11836  
 QY 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGAAGTGAAGGCAATTGATGCTTGTGCTTGC 540  
 Db 11837 AAGATTGGCAGCTGTGCTTAAGGCAAGTGAAGTGAAGGCAATTGATGCTTGTGCTTGC 11896  
 QY 541 TTGATTATTTAATGATGATGTTAAAGTGTGTTTTTTTGAATGATTAATGCAATACC 600

Db 11897 TTGATTATTTAATGATGATGTTAAAGTGTGTTTTTTTGAATGATTAATGCAATACC 11956  
 QY 601 ATCACCATCATCAATTAATAAATATGCGCTTTGGTGGCTAAGTATGCTATTTGTTGT 660  
 Db 11957 ATCACCATCATCAATTAATAAATATGCGCTTTGGTGGCTAAGTATGCTATTTGTTGT 12016  
 QY 661 GATTAATGATGATTTGGGGCAAGCTGTGACGTGTGCTAAGATGCTTGGCTAAT 720  
 Db 12017 GATTAATGATGATTTGGGGCAAGCTGTGACGTGTGCTAAGATGCTTGGCTAAT 12076  
 QY 721 GGCATGCTTGTGATGATGATGCGGCTTTGGCAATTTTCAAAAATACGAATTTTGTGC 780  
 Db 12077 GGCATGCTTGTGATGATGATGCGGCTTTGGCAATTTTCAAAAATACGAATTTTGTGC 12136  
 QY 781 CAATTAATCCACCATGCGATGCGATCAAGATGATCAAGCGAGATTTAAATTTGT 840  
 Db 12137 CAATTAATCCACCATGCGATGCGATCAAGATGATCAAGCGAGATTTAAATTTGT 12196  
 QY 841 TGGCCCTTGAGTGGCCAAATTTGAGATGCTCAAGCTGAAACCTGACGCTCAAGCAC 900  
 Db 12197 TGGCCCTTGAGTGGCCAAATTTGAGATGCTCAAGCTGAAACCTGACGCTCAAGCAC 12256  
 QY 901 CACCAATCCATATTTTGTGATGCTTAATAATCAAGTGCAGGCGTCCAAATTTTACCGCC 960  
 Db 12257 CACCAATCCATATTTTGTGATGCTTAATAATCAAGTGCAGGCGTCCAAATTTTACCGCC 12316  
 QY 961 GACACCAACATATGCTGATGCTTTTGGCATGCTGCTTCAATATGCTGAACAGTGT 1020  
 Db 12317 GACACCAACATATGCTGATGCTTTTGGCATGCTGCTTCAATATGCTGAACAGTGT 12376  
 QY 1021 TTGCGATTGAGCTGATGATGCGGATGATGCTTGTGCTTGAATGATGCTG 1080  
 Db 12377 TTGCGATTGAGCTGATGATGCGGATGATGCTTGTGCTTGAATGATGCTG 12436  
 QY 1081 TGCCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 Db 12437 TGCCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12496  
 QY 1141 GGGGGGCAATATCCTGGCTTCTTGGTTAATGCCGGGCTGATGATGATGATGATGATG 1200  
 Db 12497 GGGGGGCAATATCCTGGCTTCTTGGTTAATGCCGGGCTGATGATGATGATGATGATG 12556  
 QY 1201 TGAAGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
 Db 12557 TGAAGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12616  
 QY 1261 ATGCGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
 Db 12617 ATGCGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12676  
 QY 1321 CAAAAAATTCACAGCTGACAGTGCAGAGTGCACCAACGACGACGATATTTTGTGA 1380  
 Db 12677 CAAAAAATTCACAGCTGACAGTGCAGAGTGCACCAACGACGACGATATTTTGTGA 12736  
 Db 12677 CAAAAAATTCACAGCTGACAGTGCAGAGTGCACCAACGACGACGATATTTTGTGA 12736  
 RESULT 3  
 ADL02987/c  
 ID ADL02987 standard; DNA; 1410 BP.  
 XX ADL02987;  
 XX  
 XX 06-MAY-2004 (first entry)  
 XX  
 XX DNA encoding a M. catarrhalis protein #673.  
 XX  
 XX de; gene; Moraxella catarrhalis; infection.  
 XX  
 XX Moraxella catarrhalis.  
 XX  
 XX OS  
 XX PN US6673910-B1.  
 XX  
 XX 06-JAN-2004.



XX 04-APR-2000; 2000US-00540236.  
 PF 08-APR-1999; 99US-0128416P.  
 PR (GENO-) GENOME THERAPEUTICS CORP.  
 XX Breton GL;  
 PI MPI; 2004-178127/17.  
 DR P-PSDB; ADL04907.  
 XX  
 PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for  
 PT preparing a composition for diagnosing, preventing or treating infection  
 PT caused by Moraxella catarrhalis.  
 PS Disclosure; SEQ ID NO 673; 349pp; English.  
 CC The invention relates to an isolated nucleic acid encoding an Moraxella  
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a  
 CC composition for diagnosing, preventing or treating infection caused by  
 CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.  
 CC catarrhalis protein.  
 XX  
 SQ Sequence 1410 BP; 409 A; 307 C; 302 G; 392 T; 0 U; 0 Other;  
 Query Match 99.3%; Score 1370.4; DB 12; Length 1410;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1374; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TAAAGGTGACCACTGTTAAACAACCTTATCAGCTGCTCATTTAATTTTAATGATC 60  
 DB 1392 TAAAGGTGACCACTGTTAAACAACCTTATCAGCTGCTCATTTAATTTTAATGATC 1333  
 QY 61 CATGTAGACCAAGAGTGTACAAAAGCAAGCCCTAGGGGCTTGACTGGCTGCCAA 120  
 DB 1332 CATGTAGACCAAGAGTGTACAAAAGCAAGCCCTAGGGGCTTGACTGGCTGCCAA 1273  
 QY 121 TACTGCCGATCAAGTGTACCAACTGATATATTTTATGATGCAATGTCATCACTTT 180  
 DB 1272 TACTGCCGATCAAGTGTACCAACTGATATATTTTATGATGCAATGTCATCACTTT 1213  
 QY 181 GAGTAAATCTTGCTCAATTAATGCGGCACTTGAACGATTAATACACTATCCACCA 240  
 DB 1212 GAGTAAATCTTGCTCAATTAATGCGGCACTTGAACGATTAATACACTATCCACCA 1153  
 QY 241 CTGCATTAATGCTGAGAGCTCACTAAATCTTGAACCTTACCAAGGCCCAACCAAT 300  
 DB 1152 CTGCATTAATGCTGAGAGCTCACTAAATCTTGAACCTTACCAAGGCCCAACCAAT 1093  
 QY 301 CAATGCCAGTATGATATGCAATTAACCGTGCACAGCCCATCATATGCTGCATGTGA 360  
 DB 1092 CAATGCCAGTATGATATGCAATTAACCGTGCACAGCCCATCATATGCTGCATGTGA 1033  
 QY 361 GCCGATGTTTGTACCTTTTGAATCATTAATATGCTTGTATCATATCATATATA 420  
 DB 1032 GCCGATGTTTGTACCTTTTGAATCATTAATATGCTTGTATCATATCATATATA 973  
 QY 421 TTCACAGCATGCGGAGGCTTTGAAAGCTTAAAGGATTAAGAGATGAGCATCCATGG 480  
 DB 972 TTCACAGCATGCGGAGGCTTTGAAAGCTTAAAGGATTAAGAGATGAGCATCCATGG 913  
 QY 481 AAGATTGGAGCTGTGCTTAAGGCAAGTGCAGATTAAGGATTAAGGATTAAGGATTA 540  
 DB 912 AAGATTGGAGCTGTGCTTAAGGCAAGTGCAGATTAAGGATTAAGGATTAAGGATTA 853  
 QY 541 TTTGATTTTAAATGATGATGATGATTAAGGATTAAGGATTAAGGATTAAGGATTA 600  
 DB 852 TTTGATTTTAAATGATGATGATGATTAAGGATTAAGGATTAAGGATTAAGGATTA 793  
 QY 601 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660  
 DB 792 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 733

QY 661 GATTAATGATGATTTGGGCAACGCTTGCAGTGTGTCAGCATTTGCTGGCTAAAT 720  
 DB 732 GATTAATGATGATTTGGGCAACGCTTGCAGTGTGTCAGCATTTGCTGGCTAAAT 673  
 QY 721 GGCATGCTCTTGTATGATATGCGGCTTTGGCAATTTCAAAAATACGCAATTTTGTG 780  
 DB 672 GGCATGCTCTTGTATGATATGCGGCTTTGGCAATTTTCAAAAATACGCAATTTTGTG 613  
 QY 781 CAAATTAATCCATGATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAAATTTT 840  
 DB 612 CAAATTAATCCATGATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAAATTTT 553  
 QY 841 TGGCCCTTGAAGTGCCTCAATTTGATATGTCATCAAGCTGTAATTTGACGCTCAAG 900  
 DB 552 TGGCCCTTGAAGTGCCTCAATTTGATATGTCATCAAGCTGTAATTTGACGCTCAAG 493  
 QY 901 CACCAATTCATATTTTGTATGATGATTAATTAATTAATTAATTAATTAATTAATTA 960  
 DB 492 CACCAATTCATATTTTGTATGATGATTAATTAATTAATTAATTAATTAATTAATTA 433  
 QY 961 GACACCAATTCATATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAAATTTT 1020  
 DB 432 GACACCAATTCATATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAAATTTT 373  
 QY 1021 TTTGCAATTTGAGCTGTGATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAA 1080  
 DB 372 TTTGCAATTTGAGCTGTGATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAA 313  
 QY 1081 TGTCTTGAATGATCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140  
 DB 312 TGTCTTGAATGATCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 253  
 QY 1141 GCGGCGCAATTAATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAAATTTT 1200  
 DB 252 GCGGCGCAATTAATGCGGCTTGCAGTGTGTCAGCATTTGCTGGCTAAATTTT 193  
 QY 1201 TGAAGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 DB 192 TGAAGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133  
 QY 1261 ATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 132 ATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73  
 QY 1321 CAAAAAATTCAGAGCTGACAGTCCAGAGCTGCCAAACGAGAGGCAATTTTGA 1380  
 DB 72 CAAAAAATTCAGAGCTGACAGTCCAGAGCTGCCAAACGAGAGGCAATTTTGA 13  
 RESULT 4  
 ACA21070/c  
 ID ACA21070 standard; DNA; 1344 BP.  
 XX ACA21070;  
 AC 19-JUN-2003 (first entry)  
 XX  
 DT  
 XX  
 DE Prokaryotic essential gene #2727.  
 XX  
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX  
 OS Acinetobacter baumannii.  
 XX  
 PN W020027183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002MO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.





Db 80 TCATGACGAAATTTCACGAGAAACCTGATATTCCTCAAGCCTGCTACACTTTT 24

RESULT 5  
ADA29995/C  
ID ADA29995 standard, DNA; 1416 BP.  
XX  
XX ADA29995;  
AC  
XX  
XX 20-NOV-2003 (first entry)  
DT  
XX  
DE DNA encoding Acinetobacter baumannii protein #1282.  
XX  
XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;  
KM vaccine; plant biocontrol agent.  
XX  
XX Acinetobacter baumannii.  
OS  
XX US6562958-BI.  
PN  
XX 13-MAY-2003.  
PD  
XX  
XX 04-JUN-1999; 99US-00328352.  
PF  
XX 09-JUN-1998; 98US-0088701P.  
PR  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA  
XX  
XX Breton G, Bush D;  
PI  
XX  
XX WPI, 2003-576092/54.  
DR P-PSDB; ADA34121.  
XX  
XX  
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
PT for diagnosing a bacterial disease, as components of antibacterial  
PT vaccine; as targets for antibacterial drugs, or as biocontrol agents for  
PT plants.  
PT  
XX Example; SEQ ID NO 1282; 328bp; English.  
PS  
XX  
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.  
CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
CC for diagnosing a bacterial disease, as components of antibacterial  
CC vaccine; as targets for antibacterial drugs, to detect the presence of  
CC A. baumannii and other Acinetobacter species in a sample, in screening  
CC compounds for the ability to interfere with the A. baumannii life cycle  
CC or to inhibit A. baumannii infection, and as biocontrol agents for  
CC plants. The present sequence represents DNA encoding an A. baumannii  
CC protein.  
CC  
SQ Sequence 1416 BP; 417 A; 262 C; 326 G; 411 T; 0 U; 0 Other;

Query Match 16.4%; Score 226.2; DB 9; Length 1416;  
Best Local Similarity 52.1%; Pred. No. 2,4e-54;  
Matches 718; Conservative 0; Mismatches 593; Indels 66; Gaps 7;

QY 2 AAAGCGTGAACCATGTTAAACAATTATCACCTCGCTCATTTATTTTAAATGATCC 61  
DB 1409 AACGATTTGACGAGGCAACAACTGCTGACACGCTCATTTATACCTTTTAAACATATCA 1350  
QY 62 ATGCTTGCACACAGAGTGACAAAGACACAGCCCTTACCGGTCTTGTGACTGCTGCCAAT 121  
DB 1349 AAACCTTGACATCTCGTGTAT-----AGCAATACCAATCTTCAGCTTGTGTTTAC 1298  
QY 122 ACTGCCGATCAAGTATCAACCTGATATATTTTAGATGCCAAATGTGATCACTTTG 181  
DB 1297 GTTGAACAGCTTGAACAGCTTTCTTTAAGCGTGTGCTGATGTAATTTTAAAGTGGCCT- 1239  
QY 182 AGTAAATCTTGCTCAATTTACGGGAGCTTTGACCGATTAATTAATACACTATCCACCAC 241  
DB 1238 TGAATGCTTGTTCGATGACAGGCGCATCTTCAACCAATCAATACCAACTTTGGATAT 1179  
QY 242 TGCATTAACATATGTGACAGCTCACTAAATTTTGAACCTTTTACCAAGCCCAACCAAAATC 301

Db 1178 TTTTCATGAAAGACGTAAAGAGAAAATCTGTCTTTGCGCTGTCCACTTAAATA 1119  
QY 302 AATGCCA-----GTATGATATATCCATAAACCCTGCCAAGCCCATCAATTTGCTGCAATG 335  
DB 1118 AGTGCAACCTTACCTTTTTCACCTTCAATGCGACGACTTAAGCCATCAATTTGCTGCAAGT 1059  
QY 356 GTTAGCCGATGTATTGATCCTTTTGAATCATTTAAATATATGCTGTGATCAATATATCA 415  
DB 1058 GTTGCACCAACATTAATGATCCTTTAGAGTCAATTATTAACGACATCATGACAGCTTTTA 999  
QY 416 ATATATTCACAGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGATGAGATCC 475  
DB 998 ACCTACTCACAGGGGTGCTTACTTCTTTAAATTTTAAAGTTTCAAGCATTTGACTCC 939  
QY 476 ATGGAAAGATTGGCACTGTGCTTAAGGCAATGTCAGATTAAGCATTTAGTATAGTATG 535  
DB 938 ATAGGTAAACCAATTCCTTCCCAATGCTAAACAAAGCTTAAGCATTAATGATCATTTGTG 879  
QY 536 TTGCTTTGATTTTAAATGATGATGATGTAATAAGTTGTTTAAATGATGATATTAATGCC 595  
DB 878 ATACCTTGAATATATTAATCTGAGCTTTTAATTAACCGCTGTAACCAAGACAGACCA 819  
QY 596 ATACCATCAACATCATCAATAATAATAAATATGCGGTTTGTGCTAAGTAAGTATTT 655  
DB 818 AGCGTACCGTCCGATCTCTTAAACGCGCATATTTGAT-----TAAATCCGGTGGCT 767  
QY 656 GTTGTATTAATGTGTGATTTGGGGCAAGCTTTGTATGATGTGTCAAGCATTTCTTGCT 715  
DB 766 TTAACCAAAAGCTTTGATTTGATGTGTGATCTGTAACAAAGTGAACGGCTTA----- 716  
QY 716 AATTTGCAATCGTCTTGTAGATGATGCGGTTTGGCAATTTTCAAAATATACGATTT 775  
DB 715 ----AGCATCATCTCGGTTAAATACACTTTTAAAGCCCTTGGAAATATGATCTTT 660  
QY 776 TGTGCAAAATATCAACCAATGCCATGTCGCCATCAAGATATCAAGGAGATTTAA 835  
DB 659 GCTTGTATATACCCGACATATTTCCATGACGCTTAATGTGCTTCACTCATATTTAGA 600  
QY 836 ATTTGTTCCCTTGAAGCTGCCAAATTTGATATATGCTCAAGCTGAAACCTTGACAGCTCA 895  
DB 599 ACCACTGTACCTGACAGCGTTTAAGTAGAGGTGTTTCTTAACTGAAGCTTGAATTACTTA 540  
QY 896 AGCACCAACCAATTCATTTTGTGATCGTTAATTAATTCAAAGTGCAGGCGTGCATATTA 955  
DB 539 AGAGACAGTAATCTTGG---TTGATTTTAAGTAATCTTAAGCTGTGACCAAGTTG 483  
QY 956 CCGCGACACCAACATATGCTGTGATGTTTGCATGCTGCTAATATGTCGTAA 1015  
DB 482 CCGCTTACGTCAACTTTCTTACCTGATCCTTAAGCATTAAGCAATTAAGTGTAA 423  
QY 1016 GTGCTTTTGGATTTGAGCTGTGATGAGGATGATTTGTGTGTTTGTCTTGAAGTTAA 1075  
DB 422 GTATCTTTTGGCATTTGAACTTGAATGCGCAAAATGCGAATCAATGAGTGC----- 373  
QY 1076 TCTCGTCTTGAAGTATCAATGAACAATTAACATTCGCTGATCAAGAAATGCTTTGA 1135  
DB 372 -----AGGGGCAATTAATCTGATATGCCCCACACAAATTAATCTTTTA 330  
QY 1136 GCTTTGGCGGCAATTAATCTGCGCTTTCTTGATTAATGCCGGGCTGATGATGATCTGA 1195  
DB 329 GCAATATGCTGCTGAATTTCCGGTAATGTGTGCAAGGCCCTTGAGTTAAATATCTCT 270  
QY 1196 TCGGCTGACCTTAAGTTCATCTGTCGAAGTCCCAAAATGCAAAATGAAATGAAATGCGCTGGGC 1255  
DB 269 TCTGCTTGAATTAATTAATTTCTTGATTAAGCTACCAAAATCTGTTTAAACACAGCGGA 210  
QY 1256 AGTTATATCGCAAGGTATGATTTGAGATTAACCATGATATGATCTTAATTTGACCGCTGA 1315  
DB 209 ATCTATATCTGTCCGGAGGTGTGGGGCGGAATCCGTTTACAGCACTTTGTAGCTTTGT 150  
QY 1316 TTGACCAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAACCGACGCGCATAT 1372

DB 149 TCATGCAAGAAATTACAGCAAGAAACCTGTATATTCACCAAGCTGTCAACATTTT 93

RESULT 6  
ID AAS53453 standard; DNA; 1314 BP.  
XX AAS53453;  
XX AAS53453;  
XX 13-FEB-2002 (first entry)  
XX Haemophilus influenzae DNA for cellular proliferation protein #235.  
XX Anti-sense; db; prokaryotic cellular proliferation gene; antibiotic;  
XX antibacterial; drug design.  
XX Haemophilus influenzae.  
XX W0200170955-A2.  
XX 27-SEP-2001.  
XX 21-MAR-2001; 2001WO-US009180.  
XX 21-MAR-2000; 2000US-0191078P.  
XX 23-MAY-2000; 2000US-0206848P.  
XX 26-MAY-2000; 2000US-0207727P.  
XX 23-OCT-2000; 2000US-0242578P.  
XX 27-NOV-2000; 2000US-0253625P.  
XX 22-DEC-2000; 2000US-0257931P.  
XX 16-FEB-2001; 2001US-0269308P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Haselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
XX Yamamoto RT, Xu HH;  
XX WPI; 2001-611495/70.  
XX P-PSDB; A0535594.  
XX New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids.  
XX Claim 27; SEQ ID NO 7090; 511bp; English.  
XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the genes,  
XX their use in the discovery of novel antibiotics, the essential genes  
XX themselves and the encoded proteins. The prokaryotes used are *Bacteriella*  
XX *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
XX *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
XX useful for the identification of potential new targets for antibiotic  
XX development. The antisense nucleic acids can also be used to identify  
XX proteins used in proliferation, to express these proteins, and to obtain  
XX antibodies capable of binding to the expressed proteins. The proteins can  
XX be used to screen compounds in rational drug discovery programmes. The  
XX antisense nucleic acid sequence is also useful to screen for homologous  
XX nucleic acids which are required for cell proliferation in a wide variety  
XX of organisms. The present sequence encodes an essential prokaryotic  
XX cellular proliferation protein. Note: The sequence data for this patent  
XX did not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

Query Match 11.6%; Score 160.4; DB 4; Length 1314;  
Best Local Similarity 49.6%; Pred. No. 2,4e-35;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGCTGAACCACTGTAACAACTTATCATCTGCTCATTTATTTTAAATTGATC 60  
DB 1308 TAAACATTTGAGCTTAAATGCGTAATCTTCGCCGCTTTTCAAAAGAACAACTGATC 1249

QY 61 CATGTAGCAAGAGCAGGTGCAAAAGCAGACGCCCTAGCGGTGCTTGACTGGCGCAA 120  
DB 1248 GAGACTTGCAACAGCAGGCAATATATACCA-----TATCTCGCTTGCA 1201  
QY 121 TACTGCCGATCAAGTGTACCAACCTGATATATTTTATGATGCCAATGTGATCACCTTT 180  
DB 1200 TGTGGCGTAAATAATTCATCGCTTGTTCATTTGATGCAACA-----ATTACTTTG 1147  
QY 181 GAGTAATCTTGTCTCAATTTATCGGGCATTTTGACCGATTTAATACACTATTCACCCA 240  
DB 1146 CGATGAATAATTTTGCAAGCAGCGACCATCTCCACAAACAAATACAAATATGTGTGG 1087  
QY 241 CTGATACATATGCTGACAGCTCATTAATCTTGAACCTTTGCCAAGCCCAACCCAAAT 300  
DB 1086 TTGATTTAATTAATTCACCTAATTTGTGAATAATAGCCCTTTTCCTCGCTTACCAA 1027  
QY 301 CAATGCCAGTGAATATATGCAATMAACCGTGCAAGCCCATATGCTCAATGTGTGA 360  
DB 1026 CAATGCAATTTTACCTCATATA-----AAGCCAGCCAAATGCAACATGTACT 976  
QY 361 GCGATGTTTGAACCTTTTGAATCATTAATAATGCTTGTATCATATCATATATA 420  
DB 975 CCCACATTTTGTCTTATGATCATTAATCCAAAGATGCAATGCTTGAATGCACTAA 916  
QY 421 TTGACAGCGATGGGGTAGCCCTTGAAGTCTTAAGGTATCAAGCATGGCCATGGG 480  
DB 915 TTGAATAAGATGATCTAAACCTTTGAATGAAGCAAGGCGGTGCAATTTGAATCTTAAT 856  
QY 481 AAGATTGGCAGCTGTGCTTAAGCAGTGAAGTGAAGGCAATGATAGTTATGCTTGGC 540  
DB 855 AATACCTAATGCTTTGGGCCAATGCTGTGTGCAAAATGTCATTAATATATGCGACAC 796  
QY 541 TTTGATTTTAAATGATGATTTGTAATAAGTTGTTTTTATGATATATATCCATACC 600  
DB 795 AACCAATGTAGCTTCTTCAACAGTAATAACATCTTATACATTTAATAATTTGCTT 736  
QY 601 ATCACATCATCAATTAATAATAAATCTGC-CGTTTGGTGGCTAATGCTATTTGTTG 659  
DB 735 GCCATTTTCACTTTTATAGCCAAATATATCCGACTATATTTTCCGCAAAAGAAAGGATGTTT 676  
QY 660 TGATTAATGATGATTTGGGGCAAGCTGTGTGATGATGATCAAGCATTTGCTGCTAAT 719  
DB 675 CGCTTGATTTTGTCTTTTCCCAAAAGTCAGC-----644  
QY 720 TGGCATGCTTGTGATGATGATGGCGGTTTGGCAATTTTCAAAATATACCAATTTTGTG 779  
DB 643 ---TATCTTCACTTTTCAACACACACTTATGATATATGATATATGCGTATTTTGCTT 587  
QY 780 CCAATTAATTCACATGCGCATGCTGCGCATCAAGATGATCAGGGAGATTTAAATTTG 839  
DB 586 GGGGATTAATCTTTAATATCATATATGCAATGATTAATGATCTTATGACGTTGAAAGCAG 527  
QY 840 TTGCCCTTGAAGTGCCAATTTGATATATGCTCAAGCTGAATACTTGACAGCTCAAGCA 899  
DB 526 TCGCTGCGCAGCTTTTAACTATAGTTGTCTCAAGCTGAATACTTGAAAGCTCTTAGTA 467  
QY 900 CCAACCAATTCATATTTTGGATGCTTAATATTAATCAAGTCAAGCGTGCATATTTACCGC 959  
DB 466 CATTAAGTTCAACATCTTCAATC---ACATATGCAAAAGGGGATATCCCAATTTTCCGC 410  
QY 960 CGACACCAACATATGATGCTGATGCTTTTGGCATCTCGGCTATATATGCTGAACAGTGC 1019  
DB 409 CCAATACCACTTTTCAACACGACGCTTGGCCATTTTCAATTAATTAAGATTAACGATAC 350  
QY 1020 TTTTGGCATTTGAGCTGTGATGCGCATATGCTGTGTTTGTGCTTAGTATGATCTC 1079  
DB 349 TTTTACATTTTGAACCTGTATATCCCAATTTGCTTGTGCTGCGCGCA-----298  
QY 1080 GTGCTTTGAGTGTATCAATGAACATTTGAACATGCTGATCAACAGAAATGCTTGAAGCTT 1139  
DB 297 -----GAATTAATTCATATATGCGCATTTACTTCACTCCGCTTTTAA 257



|    |      |   |      |
|----|------|---|------|
| Db | 6433 | -TATCTTCACTGTTTCAACACACCTACTCTTAGCTTAAGCATTAATGATAAATGCGTAATTTTGCTT | 587  |
| Qy | 780  | CCAAATTAATCCACCAATGCCATTCGTCCGATCAAGATGATACGCGAGAGATTTAAATATG       | 839  |
| Db | 586  | GCGCATTAATCTTCTTAATTCATATAGCGATCCATATGATCTTCACTACGTTCAAGACAG        | 527  |
| Qy | 840  | TTGCCCTTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGA AAACTTGA CAGCTCAAGCA      | 899  |
| Db | 526  | TGCGTCGCGCAGCTTTTAAAGCTATATAGTGTGTCCTCAAGCTGGA AAACTTAAGACTTA       | 467  |
| Qy | 900  | CCACCAAAATCCATATTTTGGATTCGTTAATTAATTAAGACGAGCGCGCTGCCATTTAACC       | 959  |
| Db | 466  | CATAAAGTTCCAAATCTTCATTC--AAACAATGACAAAAGCGGAATCCCAATATTTCCG         | 410  |
| Qy | 960  | CGACACCAACATCATATGCTCGATGTTTGGCATCTCCGCTACTATATGTCGTAACAGTGC        | 1019 |
| Db | 409  | CCATACCAACTTTGACACACAGAGCTTTGCGCATTTCAATAAATTAAGTAAAGTAC            | 350  |
| Qy | 1020 | TTTTGGCATTTGAGCGCTGTGATGCGGATGATGTGTGTGTTGTCCTTGAAGTTAGATCTC        | 1079 |
| Db | 349  | TTTTTACCAATTTGAACCTGTAAATCCCAAAATGAGCTTGTGCGTCGCGGCA-----           | 298  |
| Qy | 1080 | GTGCTTGAGTGATCATTAAGAAACAATGAACATGCGCTGATCAAGAAATGCTTGAAGTT         | 1139 |
| Db | 297  | -----GAATATATTAATATGCGCATTAATCTTCCACTCCGCTTAA                       | 257  |
| Qy | 1140 | TGCGCGCGACATATCTCGCGCTCTTCTGGGTTAATATCGCGGCGTATGATGATCTGATCG        | 1199 |
| Db | 256  | GTGCGGTTTGAATTTCTGTGTGTTTTTACCAGAACGCTCGGGCTAATCAATCATATGCG         | 197  |
| Qy | 1200 | CTGAGCTTAAGATTCACCTGTCCAAGCTGCCAAAATGAACAATGAACGCTGTGGGCACTT        | 1259 |
| Db | 196  | TTTCAAGTAAACCATTTCCGATTTTAAACTACACAGATGAAGAAGGATATTTGAGGAAGTT       | 137  |
| Qy | 1260 | TATCTGCCAAGTATGAGATTGGGATTAACATGATGATCTTAATCTTGTGACCGTGATTTGA       | 1319 |
| Db | 136  | TATCAATATCACAGTATGAGATTTTTCGAGTATCAATCAACGAATATTTACCTGTGTGGGATA     | 77   |
| Qy | 1320 | CCAAAAATTTACAGAGCTGACAGTCCAGAGCTGCCAAACGACGACGATATTTT               | 1377 |
| Db | 76   | AGAGATATATCCACACAGAAAGACCTGTTTTCACAGCCCGATGATGTAATATTTT             | 19   |

RESULT 8

ADL46439/c

ID ADL46439 standard; DNA; 1314 BP.

XX ADL46439;

AC

XX

XX

XX

XX

20-MAY-2004 (first entry)

DE

XX

asparate semialdehyde dehydrogenase DNA #1.

KM

de; gene; antibacterial;

KM

UDP-N-acetylglucosamine 1-carboxyvinyl transferase-1;

KM

CTP: CMP-3-deoxy-D-manno-oculosonate transferase;

KM

UDP-N-acetylmuramylalanyl-D-glutamate-2-6-diaminopimelate ligase;

KM

D-alanine-D-alanine adding enzyme; D-alanine-D-alanine ligase;

KM

UDP-N-acetylpyruvylglucosamine reductase;

KM

UDP-N-acetylglucosamine pyrophosphorylase;

KM

UDP-N-acetylmuramoylalanine-D-glutamate ligase;

KM

DP-N-acetylmuramoylalanine ligase; asparate semialdehyde dehydrogenase;

KM

UDP-N-acetylmuramoylalanyl-D-glutamate; X-ray diffraction analysis.

XX

XX

OS

Haemophilus influenzae.

XX

XX

WO2003087353-A2.

PN

23-OCT-2003.

XX

XX

08-APR-2003; 2003WO-CA000481.

PF

08-APR-2002; 2002US-0370899P.

XX

XX

RR

08-APR-2002; 2002US-0370915P.  
PR 09-APR-2002; 2002US-0371107P.  
PR 09-APR-2002; 2002US-0371185P.  
PR 31-MAY-2002; 2002US-0365425P.  
PR 06-JUN-2002; 2002US-0366283P.  
PR 01-AUG-2002; 2002US-0400348P.  
PR 08-NOV-2002; 2002US-0424395P.  
PR 06-NOV-2002; 2002US-0425200P.  
PR 24-DEC-2002; 2002US-0436345P.  
PR 24-DEC-2002; 2002US-0436349P.  
PR 26-DEC-2002; 2002US-0436568P.  
PR 27-DEC-2002; 2002US-043675P.  
PR 27-DEC-2002; 2002US-0436734P.  
PR 27-DEC-2002; 2002US-0436885P.  
PR 27-DEC-2002; 2002US-0436889P.  
PR 27-DEC-2002; 2002US-0436893P.  
PR 30-DEC-2002; 2002US-0436900P.  
XX XX  
XX XX  
PA (AFPI-) AFFINUM PHARM INC.  
PI Edwards A., Dharamsi A., Vedadi M., Domagala M., Houston S., Awrey D.;  
PI Beattie B., Mansbury K., Ouyang H., Vallee F., Richards D., Nethery K.;  
PI Vitaris C., Budziszka K., Pinder B., Alam MZ, Tai M., Canadian V;  
PI Kanagaraiah D., Thalakada R;  
XX XX  
XX WPI; 2003-865361/80.  
DR P-P8DB; ADL46440.  
XX XX  
PT New recombinant bacterial enzymes involved in cell membrane biogenesis,  
PT useful for designing potential antibacterial agents.  
PS Claim 467; SEQ ID NO 157; 407pp; English.

The invention relates to isolated, recombinant polypeptides (I) that have at least one activity of specified bacterial enzymes involved in cell membrane biogenesis. (I) are: UDP-N-acetylglucosamine 1-carboxyvinyl transferase-1 of *Streptococcus pneumoniae* (S.p), *Pseudomonas aeruginosa* (P.a.) or *Staphylococcus aureus* (S.a.); CTP-CMP-3-deoxy-D-manno-oculosonate transferase of *Bacterichia coli* (B.c.) or *Haemophilus influenzae* (H.i.); UDP-N-acetylmuramylalanyl-D-glutamate-2,6-diaminopelate ligase of P.a.; D-alanine:D-alanine adding enzyme of S.a. or P.a.; D-alanine-D-alanine ligase of *Enterococcus faecalis* (E.f.); UDP-N-acetylpyruvoylglucosamine reductase of P.a. or H.i.; UDP-N-acetylglucosamin pyrophosphorylase of E.f., H.i. or S.a.; UDP-N-acetylmuramoylalanine D-glutamate ligase of E.f. or H.i.; DP-N-acetylmuramate:alanine ligase of E.C.; and aspartate semialdehyde dehydrogenase of H.i and UDP-N-acetylmuramylalanyl-D-glutamate (sic) of H.i. Crystallization (I) are used to determine (by X-ray diffraction analysis) the structural coordinates of (I), and these then used to design modulators of (I), potential therapeutic agents for treating diseases caused by (I), potential bacteria. This sequence represents a DNA of the invention.

Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

Query Match 11.6%; Score 160.4; DB 11; Length 1314;  
Best Local Similarity 49.6%; Pred. No. 2.4e-35;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7

1 TAAAGCGGAACCAAGTTAACAATATCACCCTGCGCATTAATTATTTTAAATTGCATC 60  
Db TAAACATTGACTTAATTCGGTAATTTCTCCCGCGCTTTCAAAGAAGCAACTGTATC 124  
61 CATGCTGACACAGCAGGTGACAAAGCACAGCCTAGCGGTGCTTGACTGCGTCCAAA 120  
Db GAGACTTGACAAACAGCAGCGCGCAATTAATAC-----TATCTCCGCTTTGCAA 120  
1248 GAGACTTGACAAACAGCAGCGCGCAATTAATAC-----TATCTCCGCTTTGCAA 120  
OY 121 TACTGCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGATCACTTT 180  
Db 1200 TGTTGGGCGTAAAAATTTCTATCGCTTTGTCATTGTATGCAACAA-----ATAACTTGG 114  
181 GAGTAATCTTGCTCAATTAATCGGAGGATCTTGAGCCGATTAATTAATGACATATCCACCA 240

Db 1146 CGATGAAATTTTGCAAGCAGCCACCACCTCCAGCCAAACAAATATATATGTGG 1087  
Qy 241 CTGCATACATATGTGAGAGCTCAGTAAATCTTGACCTTTACCAAGCCACCACCAAT 300  
Db 1086 TTGATTAATTAATTCAGCTAATTTCTGAAAATAGCCCTTTCCGCTCGGCTAGCA 1027  
Qy 301 CAATGCCAGTATGATATGCAATMAACCGTCCAAAGCCCATCAATTCGCAATGGTTGA 360  
Db 1026 CAATGCAATTTACCCCAATATA-----AAGCCAGCCAAATGCAACATCTGACT 976  
Qy 361 GCCGATGTTGACCTTTGATCATTAATAATGCTTGTCTATCATATCATATATA 420  
Db 975 CCCCACATTTGTGCTTTAGATCAATTAATCAACGATCCATTAAGCTGATGCACTAA 916  
Qy 421 TTCAAGCAGTGGGTAGCCCTTGAAAGTCTTAAGGGTATCAGAGATGGCATCCATGG 480  
Db 915 TTAAAAACATGATCTAACCCCTTGAATATGCAAGGTGCGGTACGAATTAATCTAAAT 856  
Qy 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAATTAAGGCATTAAGTATGCTTGGCC 540  
Db 855 AATACCATATAGCTGTGGCCAAATGCTGTGCTGCCAAATGTTCAATTAATTAAGGGACC 796  
Qy 541 TTTGATTTTAATGATGATGATGTAAGTTGTTTGTGATATATATATCCATACC 600  
Db 795 AACCAATGATGCTTCTTCACAAAGTAAATCACTTACATTTTACATTAATATGCTT 736  
Qy 601 ATCACCATCATCAATTAATTAATAATCTGC-CGTTGGTGGCTAAGTATTTGTTG 659  
Db 735 GCCATTTTCACTTTTACCCAAATATTCGACATATTTCCGCAAAAGAAACGGTATGTT 676  
Qy 660 TGATTAATGATGATTTGGGCAACGCTTGTCAAGTGTGCAAGCATTTGCTGGCTAAT 719  
Db 675 CGCTGATTTTCTGTTTCCCAAAAGTCAGCC----- 644  
Qy 720 TGGCATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAAATACGCAATTTTGG 779  
Db 643 ---TATCTTCAATGTTCACACACCTTACATTAATGATTAATGCGTAATTTTGCTT 587  
Qy 780 CCAATATATCCACATGCGCATGTCGCGCATCAAGATGATCAGAGGAGATTTAAATTTG 839  
Db 586 GGGGATATCTTTTAATCATATAGCATTCATATGATTTTATGTCACCTTCAAGACAG 527  
Qy 840 TTGCCCCCTTGAAGTGCMAATTTGAGATATGCTCAAGCTGMAAACTTGCAGCTCAAGCA 899  
Db 526 TGCTGTCGCGAGCTTTTAAGCTATAGTGTCTCAAGCTGMAAACTGAAAGCTCTAGTA 467  
Qy 900 CCAACCAATCCATTTTGGATGTTAATTAATTAATCAAGTGAGGCGCTGCCAATTTACCG 959  
Db 466 CATTAAGTTCACATCTTCATTC---AACATGACAAAGGGGAATCCCAATATTTCCGC 410  
Qy 960 CGACACCAACATCATGCTGCGCATGTTTGGCATCTCGCTACTAATGTCGTAACAGTGC 1019  
Db 409 CCAATACCACTTTCACACGACGCTTTGCGCATTTTCATTAATTAAGTATACGATAC 350  
Qy 1020 TTTTGCATTTGAAGCTGTGATGCGCATGATGTTGTGTTGTTGCTTAGTATGATCTC 1079  
Db 349 TTTTACATTTGAACCTGTAATCCCAACATTTGCTTGTGCTGCGCGCA----- 298  
Qy 1080 GTGCTTGAAGTATCAATGAACAAATTTGAACATGCTGATCAAGAAATGCTTAGGCTT 1139  
Db 297 -----GAATTAATTCATATACCGCATTAATCTTCCATCCCGCTTTAA 257  
Qy 1140 TGGCGGCAATACTCGCGGTTCTTGGGTTAATGCGGGGGCTGATGATGATCTGCGG 1199  
Db 256 GTGCGGTTGAATTTCTGAGTTTAAATCCGCAAGCCCTGGGCTAATAATATATATCCG 197  
Qy 1200 CTGAGCTTAAGAGTCACTGTCCAAAGCTGCCAAATGACAAATGAACCGCTGTGGGAGTT 1259  
Db 196 TTTCAAGTAAACATCTCTGATTTAAACTACAGTATAGAGAGGATATTTTGAAGAGTT 137  
Qy 1260 TATCTGCCAAGAGATGGGATTAACATGATGCTAATCTTGTGACCTGATTTGA 1319

Db 136 TATCATATACAGTATGATTTTTCAGATATCATACACGAATATATAGCTGTGGGATA 77  
Qy 1320 CCAAAAATATTCACAGCTGACAGTCCAGAGCTGCCAAACGAGCA CGGATATTTT 1377  
Db 76 AGGATATATCCACACAAAGAACCTGTTTTCGCAAGCCGATGATGATATATTTT 19  
RESULT 9  
ID AEC10868/c  
ID AEC10868 strand; DNA; 1314 BP.  
XX AEC10868;  
AC  
XX  
XX  
DT 20-OCT-2005 (first entry)  
XX  
DE Haemophilus influenzae MURD DNA.  
XX  
KW protein purification; antibacterial; antimicrobial; infection;  
KW drug screening; gene; ds; UDP-N-acetylmutamoylalanine-D-glutamate ligase.  
XX  
XX Haemophilus influenzae.  
PN US2005181388-A1.  
PD  
XX  
XX 18-AUG-2005.  
PF  
XX 04-OCT-2004; 2004US-00958216.  
XX  
PR 02-APR-2002; 2002US-0369511P.  
PR 04-APR-2002; 2002US-0369817P.  
PR 04-APR-2002; 2002US-0370102P.  
PR 08-APR-2002; 2002US-0370788P.  
PR 08-APR-2002; 2002US-0370929P.  
PR 08-APR-2002; 2002US-0370820P.  
PR 08-APR-2002; 2002US-0370859P.  
PR 08-APR-2002; 2002US-0370899P.  
PR 09-APR-2002; 2002US-0371067P.  
PR 09-APR-2002; 2002US-0371107P.  
PR 09-APR-2002; 2002US-0371140P.  
PR 09-APR-2002; 2002US-0371185P.  
PR 31-MAY-2002; 2002US-0385089P.  
PR 31-MAY-2002; 2002US-0385426P.  
PR 04-JUN-2002; 2002US-0385751P.  
PR 05-JUN-2002; 2002US-0386018P.  
PR 05-JUN-2002; 2002US-0386367P.  
PR 05-JUN-2002; 2002US-0386548P.  
PR 05-JUN-2002; 2002US-0386533P.  
PR 05-JUN-2002; 2002US-0386577P.  
PR 06-JUN-2002; 2002US-0386283P.  
PR 06-JUN-2002; 2002US-0386390P.  
PR 06-JUN-2002; 2002US-0386430P.  
PR 06-JUN-2002; 2002US-0386601P.  
PR 06-JUN-2002; 2002US-0386826P.  
PR 06-JUN-2002; 2002US-0386869P.  
PR 31-JUL-2002; 2002US-0399972P.  
PR 01-AUG-2002; 2002US-0400348P.  
PR 05-NOV-2002; 2002US-0424053P.  
PR 06-NOV-2002; 2002US-0424380P.  
PR 06-NOV-2002; 2002US-0424395P.  
PR 08-NOV-2002; 2002US-0425086P.  
PR 08-NOV-2002; 2002US-0425200P.  
PR 24-DEC-2002; 2002US-0436243P.  
PR 24-DEC-2002; 2002US-0436288P.  
PR 24-DEC-2002; 2002US-0436345P.  
PR 24-DEC-2002; 2002US-0436349P.  
PR 26-DEC-2002; 2002US-0436566P.  
PR 26-DEC-2002; 2002US-0436567P.  
PR 26-DEC-2002; 2002US-0436568P.  
PR 27-DEC-2002; 2002US-0436756P.  
PR 27-DEC-2002; 2002US-0436708P.  
PR 27-DEC-2002; 2002US-0436734P.



PR 27-DEC-2002; 2002US-0436804P.  
PR 27-DEC-2002; 2002US-0436834P.  
PR 27-DEC-2002; 2002US-0436842P.  
PR 27-DEC-2002; 2002US-0436861P.  
PR 27-DEC-2002; 2002US-0436885P.  
PR 27-DEC-2002; 2002US-0436889P.  
PR 27-DEC-2002; 2002US-0436933P.  
PR 27-DEC-2002; 2002US-0436900P.  
PR 30-DEC-2002; 2002US-0436947P.  
PR 30-DEC-2002; 2002US-0436971P.  
PR 30-DEC-2002; 2002US-0436987P.  
PR 30-DEC-2002; 2002US-0437013P.  
PR 30-DEC-2002; 2002US-0437038P.  
PR 30-DEC-2002; 2002US-0437141P.  
PR 31-DEC-2002; 2002US-0437281P.  
PR 31-DEC-2002; 2002US-0437527P.  
PR 31-DEC-2002; 2002US-0437620P.  
PR 31-DEC-2002; 2002US-0437638P.  
PR 02-APR-2003; 2003MO-CA000462.  
PR 04-APR-2003; 2003MO-CA000464.  
PR 08-APR-2003; 2003MO-CA000481.  
PR 08-APR-2003; 2003MO-CA000485.  
XX (AFPI-) AFFINIDM PHARM INC.  
XX  
PI Edwards A, Dharanji A, Vedadi M, Alam MZ, Arrowmuth C, Awrey DE;  
PI Beattie B, Buzadzija K, Canadien V, Domagala M, Houston S;  
PI Kanagarajah D, Li Q, Mansoury K, McDonald M, Nethery-Brook K, Ng I;  
PI Ouyang H, Pinder B, Richards D, Tai M, Thalakada R, Vallee F;  
PI Viag C;  
XX WPI; 2005-628189/64.  
XX P-PSDB; AEC10869.  
XX  
PT New composition comprising purified polypeptides from bacteria (e.g.  
PT *Escherichia coli*), useful for diagnosing, preventing or treating  
PT microbial infections, or in pharmacogenomic or drug screening procedures.  
XX  
PS Claim 65; SEQ ID NO 400; 667pp; English.  
XX  
CC The invention relates to a composition (I) comprising purified  
CC polypeptides from bacteria. Also described: (1) a crystallized,  
CC recombinant polypeptide comprising an amino acid sequence of (I), where  
CC the polypeptide is in crystal form; (2) a crystallized complex comprising  
CC the crystallized, recombinant polypeptide and a co-factor or a small  
CC organic molecule, where the complex is in crystal form; and (3) a host  
CC cell comprising a nucleic acid encoding a polypeptide of (1), where a  
CC culture of the host cell produces at least about 1 mg of the polypeptide  
CC per liter of culture and the polypeptide is at least about one-third  
CC soluble as measured by gel electrophoresis. The composition and methods  
CC are useful for diagnosing, preventing or treating diseases, such as  
CC microbial infections. These may also be used in pharmacogenomic or drug  
CC screening procedures. The present sequence represents a *Haemophilus*  
CC influenzae UDP-N-acetylmutamylalanine-D-glutamate ligase gene sequence,  
CC which is used in an example from the present invention.  
XX  
XX  
SQ Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

Query Match 11.6%; Score 160.4; DB 14; Length 1314;  
Best Local Similarity 49.6%; Pred. No. 2,4e-35;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGGTGAACATGTAAACAACCTTATCATCTGCTCATTTATTAATTTTAAATGATC 60  
DB 1308 TAAACATTTGAGCTAAAGCGTAAATCTTCGCGCGCTTTTCAAAAGAACCAATGATC 1249  
QY 61 CATGTGACACAGACAGGTGACAAAGACAGCGCTGAGCGTGTGATGCTGCGGCA 120  
DB 1248 GAGATCTTGACAGAGCGACCAATATATACCA-----TATCTCGCTTGCA 1201  
QY 121 TACTGCCGATCAAGTGTACCAACTGATATATTATTTAGATGCAAAATGTGATCATCCTTT 180  
DB 1200 TGTGGGCGTAAATAATCTATCGTGTTCATTTGATGACAA-----ATTAATTTTG 1147

QY 181 GAGTAATCTTGCTCAATTATGAGGAGATCTTGACCGATTAATACACTATCCACCA 240  
DB 1146 CGATGAAATTTTTCAGAGCAGCCACCATCTTCACAAAACATTAATATGTGTGG 1087  
QY 241 CTGCATACATATGCTGACAGCTCATATAATCTTGACCTTTACCAAGCCACCAAT 300  
DB 1086 TTGATTAATTAATTCAAGCTAATTTGAAAAATAGGCCCTTTTCGTCCTCGCTGCA 1027  
QY 301 CAATGCCAGTGTGATATGCTAATAACCGTCCAGCCCATCAATTGCTGCAATGGTTGA 360  
DB 1026 CAATGCAATTTACCTCATATATA-----AACCCAGCCCAATGACGCAACTGACT 976  
QY 361 GCGATGTTTGTACTTTGAAATCATTTAAATATGCTTGCTATCAATATCAATATA 420  
DB 975 CCCCACATTTGTTGCTTTAGATCATTAATCCAGCAATGCAATAGCTTGATGCACTTA 916  
QY 421 TTCACAGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGTATCAAGCATGCGCATCGAG 480  
DB 915 TTGAAAGCATGATCTTAACCTTTGAAATGACGAGATGGGATGCAATTTGAATCTAAAT 856  
QY 481 AAGATTGGACGCTGTGCTTAAGCAAGTGAATTAAGCATTAAGTTAGTTAGCTTGGC 540  
DB 855 AATACCTATAGCTTGTGCAATGCTGTGCTGCAAAATGTCATATTAATGAGGAGCC 796  
QY 541 TTTGATTTTAAATTGATGATTTGTAAGTTGTTTGTGATGATATATGATCCATACC 600  
DB 795 AACCAATGATGCTTTCACAGTAATAATCACTTATCTTATCAATTAATATGCTT 736  
QY 601 ATCACCATGATCAATATAATTAATAATCTGC-CGTTGTTGCGCTAAGTAAAGCTTTGTTG 659  
DB 735 GCCATTTTCAATTTTATGCAATTAATCGCACATATTTTCCGCAAAAGAAACGATGTTT 676  
QY 660 TGAATAGTGTGATTTGGGCGACGCTGTGATGCTGATGCAAGCATTGCTGCTTAAT 719  
DB 675 CGCTGATTTTTCGTTTCCCAAAAGTCAGCC----- 644  
QY 720 TGCCATGCTTGCTGATGATGATGCGGCTTTGGCAATTTTCAAAATATGCAATTTTGG 779  
DB 643 ---TATCTTCATTTGTTCAACACACTTATGACATTAATGATTAATGCGTAATTTGCTT 587  
QY 780 CCAATATATCCACCATGATCGTGCAGTCAAGATGATCAGCGGAGAGATTTAAATTTG 839  
DB 586 GGGATATCTTCTTAATCATATAGCATTCATATGATCTTCAAGTCAAGTCAAGAG 527  
QY 840 TTGCCCCCTTGAGCTGCCAATTTGAGATATGCTCAAGTGAATTTGACAGCTCAAGCA 899  
DB 526 TCGCTGCCGACGTTTAAAGCTTAAGTATAGTGTCTCAAGTGAATAAGTGAAGCTTAAGTA 467  
QY 900 CCACCAATTCATATTTTGGATGCTTAATTAATTAATCAAGTCAAGCGCTCAATTTACCG 959  
DB 466 CATTAAGTTTCAATCTTCAATTC---AACAAATGACAAACCGGAAATTCATATTTTCGG 410  
QY 960 CGACACCAACATCATGCTGCAATGTTTGGCATCTGCGCTACTAATGTCGTAACAGTGC 1019  
DB 409 CCATACCACTTTTCAACACGACGCTTTCGCAATTTTCAATTAATTAAGTATGATCGTAC 350  
QY 1020 TTTTGCATTTGAGCGCTGTGATGCGATGATTTGCTGTTGCTTGAATTTGATCTTC 1079  
DB 349 TTTTACATTTGAACTGTATATCCCAATTTGCTTGTGCTGCGCGGCA----- 298  
QY 1080 GTGCTTGAAGTATTAATGAACAAATTTGAACATCGTGTATCAAGAAATGCTTGAAGCTT 1139  
DB 297 -----GAATTAATTAATATGCGCAATTAATTCCTACCTCCGCTTAA 257  
QY 1140 TGGCGGCGCAATTAATCTCGGCTTCTTGAGGTTAATGCGGAGGCTGATGATGATGATGCG 1199  
DB 256 GTGCGGTTGATTTTGTGCTTTTATCCGCAAGCCCTGGCTTAATTAATATATATCCG 197  
QY 1200 CTGAGCTTAAGATTCATCTGTCCAAAGCTGCCAAATGACAAATGAAACGCTGTGGGAGATT 1259  
DB 196 TTTCAATTAACATCTCTGATTTAACTACAGTATGAAGAGGATATTTTGAAGAGATT 137







WP AEB39175\_34 340001 3503610

Query Match 11.4%; Score 157.6; DB 14; Length 110000;  
Best Local Similarity 51.9%; Pred. No. 1.3e-33;  
Matches 467; Conservative 0; Mismatches 409; Indels 24; Gaps 4;

QY 170 GCATCACCTTGTGATTAATTTGGTCAATTCGCGGCGATCTTGACCGATTAAATACACA 229  
DB 39768 GGAACCTACTTGGCTTAAGCCGATTCATCTTGCAGCATCTTCACCAATCAAAACATG 39827

QY 230 CTATCCACCACGATCAATACATATGTCAGAGCTCACTAAATCTTGACCTTACCGAACG 289  
DB 39828 GAACGACCAATTTCTGAATCTGCGCTGGCTTAATTTGAAATCCGCCCTTTTCCCTGT 39887

QY 290 CCAACCAAAATCAATGCGAGTATATGATATGCAATTAACCGTGCACGCCCATCAATTTGCT 349  
DB 39888 CCAACCGCAATTAAGAACATTTTCCCTTGATGAACCC---TCTATTCATTAATAGCA 39944

QY 350 GCATAGGTGAGCCGATGTTGTGATCTTTGAAATCATTAAATATGCTTGTATCAATA 409  
DB 39945 GAATGTGACCAATATTTGTTCTTGAATCAATTAATCAACCTACTCATCCACT 40004

QY 410 TCATCAATATATTCACAGGAGTGGTGGCTTTGAAAGCTTAAAGGTATGAAGATG 469  
DB 40005 TCTCTTACCACTGCGATCGATGCGATGCGATGCGAGAGGTTTAAACATTAAGAAATG 40064

QY 470 GCATCATGGGAATGTCAGCTGTGCTTAAGGCAATGCAATTAAGCATTAAGTATAGG 529  
DB 40065 TGTGATAGAGATTCGCCGTGCTTGGCTTAAGGCAACAGCAACAGCGATTCATCCA 40124

QY 530 TTATGCTTGCCTTTGATTTTAAATGATGATGTTGTTAAAGTTGTTTGTATGATAT 589  
DB 40125 TTATGACCCCTTTTAAATATGATTCACAGGCAAAAGAGCTCATTCCTTTCGCC 40184

QY 590 AATGACATACCATCACATCATTAATTAATTAATTCGCGCTTGTGCTTAAGTATG 649  
DB 40185 AATTAAGTATATCTTCTTGTAAATCAGCCCAATTTCCCAT---GAAG 40233

QY 650 CTATTTGTTGTATTAATGATGATTTGGGCAACGCTGTCACTGATGTCAGCAATTCG 709  
DB 40234 GTGCGCTTTGSCAAAGAGATGATTTATATCTGTTGACAGATTTGATGAGAAATG 40293

QY 710 TTGGCTAAATGTCATCTGTTGTGATGATGTCGCGTTTGGCAATTTTCAAAATACG 769  
DB 40294 T-----ATAGCGCTTCTTCATTAATTAATTAACAGCCTTGACACCATTAATCTT 40346

QY 770 AATTTTGTGCAATATCAACATGCGATGCGATGCGATGATGATGAGGAGAG 829  
DB 40347 TGTTTTGTGCTGATAGCTTCCATATGATGATGCTGCTCATAATGATACGCGTTAC 40406

QY 830 TTTAAATTTGTTGCCCCCTTGAGCTGCAATTTGAGATATGCTCAAGCTGAAAACTTGAC 889  
DB 40407 TTAAGATGCTGCAACCACTGCGCAAAAGATTAATTAATTCATTAAGAAATCTGAT 40466

QY 890 AGCTCAAGCAACCAATC---CATATTTGATGTTAATTAATTAAGTGAAGCGCTG 946  
DB 40467 AATTCATAATCCCAATCTGATGATGTCATCAATCCAACTATCTAAGACAGAGTG 40526

QY 947 CCAATTTACCGCGACCAACATCATGCGTCAATTTTGGCAATTCGCTTACTAAT 1006  
DB 40527 CCGATTTTTCGGCTTACAGTACGCAACCCGCGCGCTTACGATTTCTTCCACCAAA 40586

QY 1007 GTTCGTAACAGTCTTTTGGCATTTGAGCTGTGATGCGCATGATTTGATGTTGTGCT 1066  
DB 40587 GTTCGTAACAGTCTTTTGGCATTTGAGCTGTGATGCGCATGATGAGGCGCTAATCTCT 40646

RESULT 13  
AEB42401\_28  
Continuation (29 of 34) of AEB42401 from base 280001 (L. pneumonia DNA SEQ ID NO 6733  
WP Sequence split into 34 fragments LOCUS AEB42401 Accession Aeb42401  
WP Fragment Name Begin End  
WP AEB42401\_00 1 110000

WP AEB42401\_01 100001 210000  
WP AEB42401\_02 200001 310000  
WP AEB42401\_03 300001 410000  
WP AEB42401\_04 400001 510000  
WP AEB42401\_05 500001 610000  
WP AEB42401\_06 600001 710000  
WP AEB42401\_07 700001 810000  
WP AEB42401\_08 800001 910000  
WP AEB42401\_09 900001 1010000  
WP AEB42401\_10 1000001 1110000  
WP AEB42401\_11 1100001 1210000  
WP AEB42401\_12 1200001 1310000  
WP AEB42401\_13 1300001 1410000  
WP AEB42401\_14 1400001 1510000  
WP AEB42401\_15 1500001 1610000  
WP AEB42401\_16 1600001 1710000  
WP AEB42401\_17 1700001 1810000  
WP AEB42401\_18 1800001 1910000  
WP AEB42401\_19 1900001 2010000  
WP AEB42401\_20 2000001 2110000  
WP AEB42401\_21 2100001 2210000  
WP AEB42401\_22 2200001 2310000  
WP AEB42401\_23 2300001 2410000  
WP AEB42401\_24 2400001 2510000  
WP AEB42401\_25 2500001 2610000  
WP AEB42401\_26 2600001 2710000  
WP AEB42401\_27 2700001 2810000  
WP AEB42401\_28 2800001 2910000  
WP AEB42401\_29 2900001 3010000  
WP AEB42401\_30 3000001 3110000  
WP AEB42401\_31 3100001 3210000  
WP AEB42401\_32 3200001 3310000  
WP AEB42401\_33 3300001 345687

Query Match 11.4%; Score 157.6; DB 14; Length 110000;  
Best Local Similarity 51.9%; Pred. No. 1.3e-33;  
Matches 467; Conservative 0; Mismatches 409; Indels 24; Gaps 4;

QY 170 GCATCACCTTGTGATTAATTTGGTCAATTCGCGGCGATCTTGACCGATTAAATACACA 229  
DB 96841 GGAACCTACTTGGCTTAAGCCGATTCATCTTGCAGCATCTTCACCAATCAAAACATG 96900

QY 230 CTATCCACCACGATCAATACATATGTCAGAGCTCACTAAATCTTGACCTTTACCAACG 289  
DB 96901 GAACGACCAATTTCTGAATCTGCGCTAATTTCTGGAATTCGCCCTTTTCCCTGT 96960

QY 290 CCAACCAAAATCAATGCGAGTATATGATATGCAATTAACCGTGCACGCCCATCAATTTGCT 349  
DB 96961 CCAACCGCAATTAAGAACATTTTCCCTTGATGAACCC---TCTATTCATTAATAGCA 97017

QY 350 GCATAGGTGAGCCGATGTTGTGATCTTTGAAATCATTAAATATGCTTGTATCAATA 409  
DB 97018 GAGATTTGACACCAATATTTGTTCTTGCAGATCTTAATCAACCTACTCATCTCACT 97077

QY 410 TCATCAATATATTCACAGGAGTGGTGGCTTTGAAAGTCTTAAGGATCAAGCATG 469  
DB 97078 TCTCTTACCACTGACATGATGATGCGATGCGAGGAGGTTTAAACATTAAGAAATG 97137

QY 470 GCATCATGGGAATGTCAGCTGTGCTTAAGGCAATGCGAGATTAAGCATTAAGTATAGG 529  
DB 97138 TGTGATAGAAATTTCCCGCTTCCGCTTAAGCAACAGCAAGCGCAAGCGATTCATCCA 97197

QY 530 TTATGCTTGCCTTTGATTTTAAATGATGATTTGTTAAGTATGTTTGTGATGATAT 589  
DB 97198 TTATGACCCCTTTTATTAATGATTCATCTGCAAAAGCGCTCATCTTCTTTCGCC 97257

QY 590 AATGCAATACATCAACATCAATCAATTAATTAATTAATTCGCGTGTGCTAAGTAA 649  
DB 97258 AATTAATATGATTTCTTGTTCATACAGCCCAATTTCCCAT---GAAG 97306

QY 650 CTATTTGTTGTATTAATGATGATTTGGGCAACGCTTGTATGATGCTGAAGCATTTGC 709  
DB 97307 GCGCGCTTTTGCACAAAGAGATGATTTTAATATCTGCTTGAACGATTTGGTGAAGA---- 97362

QY 710 TTGGCTAAATGGCATCGTCTTGATGATGATGCGGTTTGGCAATTTTCAAAAATACGC 769  
Db 97363 ---ACAGTGAAGACGTCTTACCATTTAAATAAACAGCCTTGGACCAACATTAATCTT 97419  
QY 770 AATTTTGGCCCAATTAATCCACCATGTCATCGTCCGATCAAGATGATCAGCGAGAGA 829  
Db 97420 TGTTCCTGCTGCTATAGGCTTCCA TAGTGATGCTGCTGCAAAATGATCAGGGGTTACA 97479  
QY 830 TTTAAATTTGTGCCCCCTTGAAGCTGCGCAATTTGAGTATGCTCAAGCTGAAAATTGAC 889  
Db 97480 TTAAGAAATCGTCGAACACATCGGCGCAAAAGAAATTAATTAATCAAGTAAAACTGGAT 97539  
QY 890 AGCTCAAGCACCAACCAATC---CATATTTGATGCTGTAATTAATTCAGTGCAGGCGTG 946  
Db 97540 AATCCCAATCCCAAAATGCTAGTATGTTTCATCATCAACATATTAAGACAGAGTG 97599  
QY 947 CCAATATTAACCGCCGACCAACCAATCATGCTGCATGTTTGGCATCTGCGCTACTAAT 1006  
Db 97600 CCGATATTTCCGGCTACAGCTACACGAAACCCGGCGCCTTAGCATTTCTCCACCAAA 97659  
QY 1007 GTGCTAACGCTGCTTTTGGCATTTGAGCTGTGATGCGCATGATTGGTGTGTTTGTCT 1066  
Db 97660 GTGCTTAAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCT 97719

## RESULT 14

ID ADT05533 standard; DNA; 14324 BP.  
ADT05533;

XX 02-DEC-2004 (first entry)  
XX Haemophilus influenzae (NTH1) contig DNA sequence - SEQ ID 569.  
XX middle ear bacterial infection; nasopharynx bacterial infection; ds;  
XX contig.  
XX Haemophilus influenzae.  
XX WO2004078949-A2.  
XX 16-SEP-2004.  
XX 05-MAR-2004; 2004WO-US007001.  
XX 06-MAR-2003; 2003US-0453134P.  
XX (CHIL-) CHILDRENS HOSPITAL INC.  
XX Bakaletz LO, Munson RS, Dyer DW;  
XX WPI; 2004-662422/64.  
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,  
XX useful for treating or preventing NTH1 bacterial infections of the middle  
XX ear and/or nasopharynx.  
XX Example 1; SEQ ID NO 569; 88bp; English.  
XX The invention comprises nucleotide sequences (genes) from the genome of a  
XX nontypeable strain of Haemophilus influenzae (NTH1). The NTH1 DNA  
XX sequences of the invention are useful for treating or preventing NTH1  
XX bacterial infections of the middle ear and/or nasopharynx. The present  
XX nucleic acid represents an NTH1 contig sequence of the invention.  
XX Sequence 14324 BP; 4572 A; 3171 C; 2271 G; 4310 T; 0 U; 0 Other;

Query Match 11.2%; Score 154.4; DB 13; Length 14324;  
Best Local Similarity 49.0%; Pred. No. 4.2e-33;  
Matches 591; Conservative 0; Mismatches 546; Indels 69; Gaps 4;

QY 172 ATCACCCTTGAAGTAATCTTGTCAATTAATCGGGGACCTTTGACCGATTAATACACACT 231  
Db 5280 ATAGCTTTGTGATGAATAATTTTGCAGAGCTGACACATCTCGACCAAAACATAACAAAT 5339  
QY 232 ATCCACCACTGATTAACATATGCTGACAGCTCATTAATCTTGAACCTTTACCAAGCCC 291  
Db 5340 AATGTGTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5399  
QY 292 ACCCAAAATCAATGCCAGTATATATGCAATTAACCGTGCCCAAGCCCATTAATGTCTGC 351  
Db 5400 GCCTACCAACAAATGCAAT-----TTAACCTCAATTAATTAATTAATTAATTAATTAAT 5450  
QY 352 AATGTGAGCCGATGTTTGAATCTTGAATCAATTAATTAATTAATTAATTAATTAATTAAT 411  
Db 5451 AACTGATCTCCCACTTTGTTGCTTTAGATTAATTAATTAATTAATTAATTAATTAATTAAT 5510  
QY 412 ATCAATATATTCACAGCATGCGGTAGCCCTTTGAAAGCTTAAGGGATTCACAGATGCG 471  
Db 5511 ATGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5570  
QY 472 ATCATAGGAAGATTGGCAGCTGTGCTTAAGCGAAGTGCATTAAGCATTAAGTATGATT 531  
Db 5571 ATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5630  
QY 532 ATGCTTGCCTTGAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 591  
Db 5631 ATGGCAGCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5690  
QY 592 TGCATATACATCATCATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 651  
Db 5691 GATATGCTTGCATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 5750  
QY 652 ATTTGTTGATTAATGCTGTGATTTGGGCAAGCTTGCAAGTGTGTCAGACATTTGCTT 711  
Db 5751 GGTTCGTCGCTGCTGATTTTCACTTCAACCAAAAGTCAGC----- 5790  
QY 712 GGCTAATTTGGATGCTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771  
Db 5791 -----TTATCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5840  
QY 772 TTTTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 831  
Db 5841 TTTTGTGCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5900  
QY 832 TAAATTTGTCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891  
Db 5901 CAAACAGTCGACGCGAGCTTTTAAGCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5960  
QY 892 CTCAGACACCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 951  
Db 5961 CTCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6017  
QY 952 ATTACCGCGACACCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1011  
Db 6018 ATTTCCGCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6077  
QY 1012 AACAGTCTTTTGGCATTTGAGCTGTGATGCGATTAATTAATTAATTAATTAATTAATTAAT 1071  
Db 6078 AACAGTACTTTTTCATTTGAACCTGTATCCCAACATTTGCTTGTGCTGCGCGG--- 6134  
QY 1072 TAGATCTGCTGCTTGAAGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1131  
Db 6135 -----GCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6170  
QY 1132 TTGAGCTTTGGCGGACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1191  
Db 6171 CGCTTAATGCGGTTTGAATTTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 6230  
QY 1192 CTGATGCGCTGAGCTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1251  
Db 6231 CATATCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6290  
QY 1252 GGGCAGTTTATTCGCAAGTATGATTTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1311

Db 6291 AGGAAATTATCAATACAGTAGGCTTTTTCGGGATCATCAACGAATATTAGCTTG 6350  
 QY 1312 GTGATTGACCAAAAATTTCAGCTGACAGCTGCGCAAAACCGACGCGCAT 1371  
 Db 6351 TTGGATTAAGAGATTAATCCACACAGAAAGACCTGTTTCCAGCCCAATGATCTGAT 6410  
 QY 1372 TTTTTC 1377  
 Db 6411 ATTTT 6416

RESULT 15  
 ADT05649/c  
 ID ADT05649 standard; DNA; 349980 BP.  
 AC ADT05649;  
 XX  
 XX 02-DEC-2004 (first entry)  
 XX  
 XX Haemophilus influenzae (NTH1) DNA sequence - SEQ ID 685.  
 DE  
 XX middle ear bacterial infection; nasopharynx bacterial infection; ds.  
 KM  
 XX Haemophilus influenzae.  
 OS  
 WO2004078949-A2.  
 PN  
 XX 16-SEP-2004.  
 PD  
 XX 05-MAR-2004; 2004MO-US007001.  
 XX  
 XX 06-MAR-2003; 2003US-0453134P.  
 PR  
 XX (CHIL-) CHILDRENS HOSPITAL INC.  
 XX PA  
 XX Bakalecz LO, Munson RS, Dyer DW;  
 XX PI  
 XX WPI; 2004-662422/64.  
 DR  
 XX  
 XX New polynucleotides of nontypeable strain of Haemophilus influenzae,  
 PT useful for treating or preventing NTH1 bacterial infections of the middle  
 PT ear and/or nasopharynx.  
 PT  
 XX Claim 1; SEQ ID NO 685; 88bp; English.  
 PS  
 CC The invention comprises nucleotide sequences (genes) from the genome of a  
 CC nontypeable strain of Haemophilus influenzae (NTH1). The NTH1 DNA  
 CC sequences of the invention are useful for treating or preventing NTH1  
 CC bacterial infections of the middle ear and/or nasopharynx. The present  
 CC nucleic acid represents an NTH1 DNA sequence of the invention.  
 CC  
 SQ Sequence 349980 BP; 105127 A; 63538 C; 70035 G; 111276 T; 0 U; 4 Other;

Query Match 11.2%; Score 154.4; DB 13; Length 349980;  
 Best Local Similarity 49.0%; Pred. No. 1.9e-32;  
 Matches 591; Conservative 0; Mismatches 546; Indels 69; Gaps 4;

QY 1172 ATACCTTGAAGTAATCTTGCTCAATTATCGGGGATCTTGACCGATTAAATACACT 231  
 Db 107917 ATAGCTTTGTGATGAAAAATTTTGCAAGCTGCACACATCTCGACAAAAACATTAACAAAT 107858  
 QY 232 ATCCACCCAGCTGATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCC 291  
 Db 107857 ATATGTGTGTTGATTAATTAATTCAGCTAATCTGAAAAATACAGCCCTTTCCGCTCC 107798  
 QY 292 ACCCAAAATCAATGCGAGTATGATGCGATTAACCGTCCAGCCCACTCAATGCTGC 351  
 Db 107797 GCTTAGCAACAATGCAAT-----TTACCCCTCAATATTAAGCCCGCTAATGCAAC 107747  
 QY 352 AATGTTGAGCGGATTTGTAACCTTTGATCATTAATAATATGCTTGTATTCATATC 411  
 Db 107746 AACTGATCTCCCACTTTGTTGCTTAAGTCAATTAATCAACGAATACATTAAGCTTG 107687

QY 412 ATCAATATATTCACAGCGATGCGGTAGCCCTTGAAGCTTAAAGGTATCAAGATGCG 471  
 Db 107686 ATGCACTAATTAAGAAACGATGATCTTAACCTTTGAAATGACAAAGCGGTACGATTTGA 107627  
 QY 472 ATTCATGGAAGATTGGCAGCTGTGCTTAAGCAAGTGCAGATTAAGGCAATGATGATGTT 531  
 Db 107626 ATCTAATTAATTAACGACAGCTGTGCAATCTGTGCTGCCAAATGTCATATTAAT 107567  
 QY 532 ATGCTTGCCTTGAATTTTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 591  
 Db 107566 ATGGCAGCCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 107507  
 QY 592 TGCATATCATCAACATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 651  
 Db 107506 GATATGCTTGCATTTGAGTTTGAAGCAATATCCGACATATTTCCGCAAAAGAAAC 107447  
 QY 652 ATTTGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711  
 Db 107446 GGTGTTGCTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 107407  
 QY 712 GGTAAATTTGGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 771  
 Db 107406 -----TTATCTCATTAATTCAAAATCTCCAGCTCAGCATTAATGATGATGATGAT 107357  
 QY 772 TTTTGTGCAATTAATTCACCATGCGATGCGGATCAATGATGATGATGATGATGATGAT 831  
 Db 107356 TTTGCTTGGCATTAATCTTAAATTCATTAATGATGATGATGATGATGATGATGATGAT 107297  
 QY 832 TAAATTTGTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891  
 Db 107296 CAAAACATGCGAGCGCGAGCTTTTAAGCTAATGATGATGATGATGATGATGATGAT 107237  
 QY 892 CTCAGCAGCACCAATCAATTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 951  
 Db 107236 CTCTGATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 107180  
 QY 952 ATTACCGCGAGCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011  
 Db 107179 ATTTCCGCGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 107120  
 QY 1012 AACAGTCTTTTGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071  
 Db 107119 AACAGTCTTTTCCATTTGAATGATGATGATGATGATGATGATGATGATGATGAT 107063  
 QY 1072 TAGATCTGCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131  
 Db 107062 -----GCAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 107027  
 QY 1132 TTGACCTTTGGCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191  
 Db 107026 CGCTTAAGTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 106967  
 QY 1192 CTGATCGCTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251  
 Db 106966 CATATCGCTTTCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 106907  
 QY 1252 GGGAGTTATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311  
 Db 106906 AGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 106847  
 QY 1312 GTGATTGACCAAAAATTTCAGCTGACAGCTGCGCAAAACCGACGCGCAT 1371  
 Db 106846 TTGGATTAAGATTAATCAACAGAAAGACCTGTTTGCAGAGCCCAATGATGAT 106787  
 QY 1372 TTTTTC 1377  
 Db 106786 ATTTT 106781

Search completed: May 12, 2006, 07:54:32  
 Job time : 913 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 12, 2006, 07:39:28 / Search time 7220 Seconds  
(without alignments)  
10864.815 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380  
Sequence: 1 taaagcgtgacacatgttaa.....acgagcgcatatttttga 1380

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID         | Description        |
|------------|--------|-------------|--------|---------------|--------------------|
| 1          | 1380   | 100.0       | 96109  | 6 AR408756    | AR408756 Sequence  |
| 2          | 1380   | 100.0       | 96109  | 6 AX067460    | AX067460 Sequence  |
| 3          | 1370.4 | 99.3        | 1410   | 6 AR450009    | AR450009 Sequence  |
| 4          | 266.8  | 19.3        | 110000 | 1 CP000082_25 | Continuation (26 o |
| 5          | 242    | 17.5        | 110000 | 1 CR543861_02 | Continuation (3 of |
| 6          | 229.6  | 16.6        | 110000 | 1 CP000082_24 | Continuation (25 o |
| 7          | 226.2  | 16.4        | 1416   | 6 AR318733    | AR318733 Sequence  |
| 8          | 160.4  | 11.6        | 1314   | 6 AX932204    | AX932204 Sequence  |
| 9          | 160.4  | 11.6        | 11498  | 1 U32793      | U32793 Haemophilus |
| 10         | 160.4  | 11.6        | 11498  | 6 AX191762    | AX191762 Sequence  |
| 11         | 160.4  | 11.6        | 110000 | 6 BD426631_11 | Continuation (12 o |
| 12         | 160.4  | 11.6        | 110000 | 6 AR274513_11 | Continuation (13 o |
| 13         | 160.4  | 11.6        | 110000 | 6 AR274513_12 | Continuation (13 o |
| 14         | 160.4  | 11.6        | 110000 | 6 AR632719_11 | Continuation (12 o |
| 15         | 160.4  | 11.6        | 110000 | 6 AR632719_12 | Continuation (13 o |
| 16         | 157.6  | 11.4        | 110000 | 1 CR628336    | Continuation (31 o |
| 17         | 157.6  | 11.4        | 110000 | 1 CR628337_28 | Continuation (29 o |

|   |    |       |      |        |               |                    |
|---|----|-------|------|--------|---------------|--------------------|
| C | 19 | 154.8 | 11.2 | 110000 | 1 AE016828_01 | Continuation (2 of |
| C | 20 | 154.4 | 11.2 | 14324  | 6 CO872956    | CO872956 Sequence  |
| C | 21 | 154.4 | 11.2 | 110000 | 1 CP000057_12 | Continuation (13 o |
| C | 22 | 154.4 | 11.2 | 349980 | 6 CO873072    | CO873072 Sequence  |
| C | 23 | 152.8 | 11.1 | 110000 | 1 AE017354_29 | Continuation (30 o |
| C | 24 | 152.4 | 11.0 | 1314   | 6 AX932206    | Continuation (5 of |
| C | 25 | 146.2 | 10.6 | 110000 | 1 BA000031_04 | Continuation (17 o |
| C | 26 | 142.2 | 10.3 | 12322  | 1 AE006048    | AE006048 Pasteurel |
| C | 27 | 139.6 | 10.1 | 1317   | 6 AX189058    | AX189058 Sequence  |
| C | 28 | 139.6 | 10.1 | 2608   | 1 ECMORX      | X1584 Escherichia  |
| C | 29 | 139.6 | 10.1 | 28277  | 1 ECMORX      | X55034 E. coli 2 m |
| C | 30 | 139.6 | 10.1 | 28277  | 6 AX191720    | AX191720 Sequence  |
| C | 31 | 139.6 | 10.1 | 110000 | 1 U00096_00   | U00096 Escherichia |
| C | 32 | 139.6 | 10.1 | 111408 | 1 EC0110K     | D10483 Escherichia |
| C | 33 | 139.6 | 10.1 | 300409 | 1 AE016755    | AE016755 Escherich |
| C | 34 | 137.4 | 10.0 | 110000 | 1 AE017340_04 | Continuation (5 of |
| C | 35 | 134.8 | 9.8  | 110000 | 1 AE005174_00 | AE005174 Escherich |
| C | 36 | 134.8 | 9.8  | 110000 | 1 AE005174_01 | Continuation (2 of |
| C | 37 | 134.8 | 9.8  | 110000 | 1 BA000007_00 | BA000007 Escherich |
| C | 38 | 134.8 | 9.8  | 110000 | 1 BA000007_01 | Continuation (2 of |
| C | 39 | 133.2 | 9.7  | 1629   | 1 ECMORD      | X17609 Escherichia |
| C | 40 | 131.6 | 9.5  | 110000 | 1 AE005674_00 | AE005674 Shigella  |
| C | 41 | 131.6 | 9.5  | 290029 | 1 AE016978    | AE016978 Shigella  |
| C | 42 | 130.2 | 9.4  | 110000 | 1 AE016827_16 | Continuation (17 o |
| C | 43 | 125.2 | 9.1  | 1323   | 6 AR377824    | AR377824 Sequence  |
| C | 44 | 124.6 | 9.0  | 110000 | 1 CP000058_46 | Continuation (47 o |
| C | 45 | 124   | 9.0  | 304558 | 1 AE017151    | AE017151 Haemophil |

## ALIGNMENTS

RESULT 1  
AR408756  
LOCUS AR408756 96109 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 35 from patent US 6632636.  
ACCESSION AR408756  
VERSION AR408756.1 GI:40159151  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 96109)  
AUTHORS Lagace,R.E., Patterson,C. and Berg,K.L.  
TITLES Nucleic acids encoding 3-ketoacyl-ACP reductase from Moraxella catarrhalis  
JOURNAL Patent: US 6632636-A 35 14-OCT-2003;  
FEATURES  
source  
1. 96109  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 1380; DB 6; Length 96109;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TAAAGCGTGAACCATGTTAAACAACCTTATCAGCTCGCTATTAATTTTAAATTGATC 60  
DB 11357 TAAAGCGTGAACCATGTTAAACAACCTTATCAGCTCGCTATTAATTTTAAATTGATC 11416  
QY 61 CATGCTAGCAACAAGCGGTGACAAAGCAAGCCCTTACGCGTCTTGAAGTGGCTGCCAA 120  
DB 11417 CATGCTAGCAACAAGCGGTGACAAAGCAAGCCCTTACGCGTCTTGAAGTGGCTGCCAA 11476  
QY 121 TACTGCCGATCAAGTGAACAACCTGATATATTTAGATGCCAAATGTGCATCAGCTTT 180  
DB 11477 TACTGCCGATCAAGTGAACAACCTGATATATTTAGATGCCAAATGTGCATCAGCTTT 11536  
QY 181 GAGTAAATCTTGCTCAATTAATCGGGGCAATCTTGACCGGATTAATACACACTATCCACCA 240  
DB 11537 GAGTAAATCTTGCTCAATTAATCGGGGCAATCTTGACCGGATTAATACACACTATCCACCA 11596

QY 241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 300  
DB 11597 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 11656  
QY 301 CAATGCGAGTATGATGCTGATTAACCGGCGCAAGCCCACTCAATGCTGATGCTGTA 360  
DB 11657 CAATGCGAGTATGATGCTGATTAACCGGCGCAAGCCCACTCAATGCTGATGCTGTA 11716  
QY 361 GCCGATGTTGTAACCTTTGATATCAATTAATATGCTGCTCAATATCATCAATATA 420  
DB 11717 GCCGATGTTGTAACCTTTGATATCAATTAATATGCTGCTCAATATCATCAATATA 11776  
QY 421 TTCAACGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGCATTCATG 480  
DB 11777 TTCAACGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGCATTCATG 11836  
QY 481 AAGATTGGGAGCTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGGCTTAATGCTTGGC 540  
DB 11837 AAGATTGGGAGCTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGGCTTAATGCTTGGC 11896  
QY 541 TTGATTTTAAATTTGATGATGATGATTAAGTATGTTTGTGATGATGATGATGATGATG 600  
DB 11897 TTGATTTTAAATTTGATGATGATGATTAAGTATGTTTGTGATGATGATGATGATGATG 11956  
QY 601 ATCAACATCATCAATTAATTAATAAATCTGCGTTGCTGCTAAGTAAAGCTATTTGTTGT 660  
DB 11957 ATCAACATCATCAATTAATTAATAAATCTGCGTTGCTGCTAAGTAAAGCTATTTGTTGT 12016  
QY 661 GATTAAGTGTGATTTTGGGCGAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 12017 GATTAAGTGTGATTTGGGCGAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 12076  
QY 721 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 12077 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12136  
QY 781 CAAATTAATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 12137 CAAATTAATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12196  
QY 841 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 12197 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12256  
QY 901 CACCAATCAATTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 12257 CACCAATCAATTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12316  
QY 961 GACACCAACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 12317 GACACCAACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12376  
QY 1021 TTGCGATTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 12377 TTGCGATTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12436  
QY 1081 TGCTTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 12437 TGCTTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12496  
QY 1141 GCGCGGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 12497 GCGCGGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12556  
QY 1201 TGAGCTTAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
DB 12557 TGAGCTTAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 12616  
QY 1261 ATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 12617 ATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12676

QY 1321 CAAAAAATTCACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 1380  
DB 12677 CAAAAAATTCACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 12736  
RESULT 2  
AX067460  
LOCUS AX067460 96109 bp DNA linear PAT 24-JAN-2001  
DEFINITION Sequence 35 from Patent WO0078968.  
ACCESSION AX067460  
VERSION AX067460.1 GI:12545080  
KEYWORDS  
SOURCE  
ORGANISM Moraxella catarhalis  
Moraxella catarhalis  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Moraxellaceae; Moraxella.  
REFERENCE  
1 Lagace, R.E., Patterson, C. and Berg, K.L.  
Nucleotide sequences of moraxella catarhalis genome  
Patent: WO 0078968-A 35 28-DEC-2000;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1..96109  
/organism="Moraxella catarhalis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:480"  
ORIGIN  
Query Match 100.0%; Score 1380; DB 6; Length 96109;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Replicated

QY 1 TAAAGCGTGAACCATGTTAAACAACTTATCACCTGCTCATTAATTTTAAATGATC 60  
DB 11357 TAAAGCGTGAACCATGTTAAACAACTTATCACCTGCTCATTAATTTTAAATGATC 11416  
QY 61 CATGCTAGCAACAAGCTGTAACAAAGCAAGCCCTGAGGCTTTGACTGCTGCCAA 120  
DB 11417 CATGCTAGCAACAAGCTGTAACAAAGCAAGCCCTGAGGCTTTGACTGCTGCCAA 11476  
QY 121 TACTGCCGATCAAGTGTACCAACCTGATATTTTGAATGCGCAATGTCACCTT 180  
DB 11477 TACTGCCGATCAAGTGTACCAACCTGATATTTTGAATGCGCAATGTCACCTT 11536  
QY 181 GAGTAATCTTCTCAATTAATGCGGCGATCTTGACGATTAATACACTATCCACCA 240  
DB 11537 GAGTAATCTTCTCAATTAATGCGGCGATCTTGACGATTAATACACTATCCACCA 11596  
QY 241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 300  
DB 11597 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 11656  
QY 301 CAATGCGAGTATGATGCTGATTAACCGTGCAGGCCCACTCAATGCTGCAATGTTTA 360  
DB 11657 CAATGCGAGTATGATGCTGATTAACCGTGCAGGCCCACTCAATGCTGCAATGTTTA 11716  
QY 361 GCCGATGTTGTAACCTTTGATATCAATTAATATGCTGCTCAATATCATCAATATA 420  
DB 11717 GCCGATGTTGTAACCTTTGATATCAATTAATATGCTGCTCAATATCATCAATATA 11776  
QY 421 TTCAACGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGCATTCATG 480  
DB 11777 TTCAACGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGCATTCATG 11836  
QY 481 AAGATTGGGAGCTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGGCTTAATGCTTGGC 540  
DB 11837 AAGATTGGGAGCTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGGCTTAATGCTTGGC 11896  
QY 541 TTGATTTTAAATTTGATGATGATGATTAAGTATGTTTGTGATGATGATGATGATGATG 600  
DB 11897 TTGATTTTAAATTTGATGATGATGATTAAGTATGTTTGTGATGATGATGATGATGATG 11956  
QY 601 ATCAACATCATCAATTAATTAATAAATCTGCGTTGCTGCTAAGTAAAGCTATTTGTTGT 660



|    |       |  |       |
|----|-------|--|-------|
| Dp | 11957 | ATGCACATCATCAATAATAATAAAAATCTGCCGTTTGGTGGCTAAGTAGCTAATTTGTGT | 12016 |
| Qy | 661   | GATAATGGTGTGATTTGGGGCAACGCTTGTCACTGTGTCACAGCATTTGCTTGGCTAAT  | 720   |
| Dp | 12017 | GATAATGGTGTGATTTGGGGCAACGCTTGTCACTGTGTCACAGCATTTGCTTGGCTAAT  | 12076 |
| Qy | 721   | GGCATTCGCTTTGGTAGATGATGGCGGTTTGGCAATTTCAAAAATAACCAATTTTGTGC  | 780   |
| Dp | 12077 | GGCATTCGCTTTGGTAGATGATGGCGGTTTGGCAATTTCAAAAATAACCAATTTTGTGC  | 12136 |
| Qy | 781   | CAAAATATCCACATGCGATCTGTCGATCAAGATGATTCAGCGGAGAGATTTAAATTTGT  | 840   |
| Dp | 12137 | CAAAATATCCACATGCGATCTGTCGATCAAGATGATTCAGCGGAGAGATTTAAATTTGT  | 12196 |
| Qy | 841   | TGCCCTTTGAGCTCCAAATTTGAGATATGCTCAAGCTGAAAATTGACAGCTCAAGCAC   | 900   |
| Dp | 12197 | TGCCCTTTGAGCTCCAAATTTGAGATATGCTCAAGCTGAAAATTGACAGCTCAAGCAC   | 12256 |
| Qy | 901   | CACCAAAATCCAAATTTTGGATCGTTAAATATCAAGTCAGGCGTCCAAATATTAACGCC  | 960   |
| Dp | 12257 | CACCAAAATCCAAATTTTGGATCGTTAAATATCAAGTCAGGCGTCCAAATATTAACGCC  | 12316 |
| Qy | 961   | GACATCCAAACATCATGCTCTGCATGTTTGGCATCTGCGCTACTAATGTGTAACAGTCT  | 1020  |
| Dp | 12317 | GACATCCAAACATCATGCTCTGCATGTTTGGCATCTGCGCTACTAATGTGTAACAGTCT  | 12376 |
| Qy | 1021  | TTTGGCAATTTGAGCTGTGATGGGAGATTTGGAGTGTGTTGCTTGGAGTTAGATCTCG   | 1080  |
| Dp | 12377 | TTTGGCAATTTGAGCTGTGATGGGAGATTTGGAGTGTGTTGCTTGGAGTTAGATCTCG   | 12436 |
| Qy | 1081  | TGCTTTGAGTGTATCAATGAACAATTTAAATCATCGCTGATCAAGAAATGCTTGAAGCTT | 1140  |
| Dp | 12437 | TGCTTTGAGTGTATCAATGAACAATTTAAATCATCGCTGATCAAGAAATGCTTGAAGCTT | 12496 |
| Qy | 1141  | GGCGGCGACAAATCACTCGGCGTTCTTGGGTTAATCCGGGGCTGATGATCTGATCGCG   | 1200  |
| Dp | 12497 | GGCGGCGACAAATCACTCGGCGTTCTTGGGTTAATCCGGGGCTGATGATCTGATCGCG   | 12556 |
| Qy | 1201  | TGAGCTTAAAGTTCACTGTTCCAAGCTGCCAAATAGCAATGAAACGCTGTGGGCAGTT   | 1260  |
| Dp | 12557 | TGAGCTTAAAGTTCACTGTTCCAAGCTGCCAAATAGCAATGAAACGCTGTGGGCAGTT   | 12616 |
| Qy | 1261  | ATCTGCCAAGGTAGAGATTGGGATTTACATGAGATGCTTAATCTTTGTACCGTGAATGAC | 1320  |
| Dp | 12617 | ATCTGCCAAGGTAGAGATTGGGATTTACATGAGATGCTTAATCTTTGTACCGTGAATGAC | 12676 |
| Qy | 1321  | CAAAAATATTCACAGCTCAAGTCCAGAGCTGCCCAACGACGACGCAATTTTTTTTGA    | 1380  |
| Dp | 12677 | CAAAAATATTCACAGCTCAAGTCCAGAGCTGCCCAACGACGACGCAATTTTTTTTGA    | 12736 |

| RESULT 3   | AR450009/c  | LOCUS                       | AR450009 | 1410 bp | DNA | linear | PAT 20-FEB-2 |
|------------|---|-----------------------------|----------|---------|-----|--------|--------------|
| DEFINITION | Sequence  | 673 from patent US 6673910. |          |         |     |        |              |
| ACCESSION  | AR450009  |                             |          |         |     |        |              |
| VERSION    | AR450009.1  | GI:42679066                 |          |         |     |        |              |
| KEYWORDS   |   |                             |          |         |     |        |              |
| SOURCE     | Unknown.  |                             |          |         |     |        |              |
| ORGANISM   | Unknown.  |                             |          |         |     |        |              |
| REFERENCE  | Unclassified.   |                             |          |         |     |        |              |
| AUTHORS    | 1 (bases 1 to 1410)   |                             |          |         |     |        |              |
| TITLE      | Breton,G.L.   |                             |          |         |     |        |              |
| JOURNAL    | Nucleic acid and amino acid sequences relating to M. catarrhalis for diagnostics and therapeutics |                             |          |         |     |        |              |
| FEATURES   | Patent: US 6673910-A 673 06-JUN-2004;   |                             |          |         |     |        |              |
| source     | Genome Therapeutics Corporation; Waltham, MA  |                             |          |         |     |        |              |
|            | location/Qualifiers   |                             |          |         |     |        |              |
|            | 1..1410   |                             |          |         |     |        |              |
|            | /organism="unknown"   |                             |          |         |     |        |              |
| ORIGIN     | /mol_type="genomic DNA"   |                             |          |         |     |        |              |

|                            |        |               |           |              |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match                | 99.3%  | Score 1370.4; | DB 6;     | Length 1410; |
| Best Local Similarity      | 99.6%; | Pred. No. 0;  |           |              |
| Matches 1374; Conservative | 0;     | Mismatches 6; | Indels 0; | Gaps 0;      |

|    |      |  |      |
|----|------|--|------|
| QY | 1    | TAAGACGGGAAACGATGTAAACAACTTATGACCTGGCTCACTTAATAATTTTAAATGTATC  | 60   |
| Db | 1392 | TAAACGGTAAACCATGTATTAACAACTTATGACCTGGCTCACTTAATAATTTTAAATGTATC | 1333 |
| QY | 61   | CATGTGACACAAAGCAGGTGA CAAAAGCACAGCCCTAGCCGGTGTGACTGGCTGCCAA    | 120  |
| Db | 1332 | CATGTGACACAAAGCAGGCGCA CAAAGCACAGCCCTAGCCGGTGTGACTGGCTGCCAA    | 1272 |
| QY | 121  | TACTGCCCGGATCAAGTGTACCACTGTATATTTTATAGATGCCAATGTGCATCACTTT     | 180  |
| Db | 1272 | TACTGCCCGGATCAAGTGTACCACTGTATATTTTATAGATGCCAATGTGCATCACTTT     | 1213 |
| QY | 181  | GAGTAAATCTTGCTCAATTTATCGGGGACCTTTGACCGATTAATATACACATATCCACCA   | 240  |
| Db | 1212 | GAGTAAATCTTGCTCAATTTATCGGGGACCTTTGACCGATTAATATACACATATCCACCA   | 1155 |
| QY | 241  | CTGCATTAACATATGTCTGAAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT     | 300  |
| Db | 1152 | CTGCATTAACATATGTCTGAAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT     | 1093 |
| QY | 301  | CAATGCCAGTATATATATGCCATTAACCGTGGCAAGCCATCAATTTGTGTGAATGTGTA    | 360  |
| Db | 1092 | CAATGCCAGTATATATATGCCATTAACCGTGGCAAGCCATCAATTTGTGTGAATGTGTA    | 1033 |
| QY | 361  | GCCGATGTTTGTACCTTTTGAATCAATTAAATATAGCTTGCTATCAATATATCATCAATTA  | 420  |
| Db | 1032 | GCCGATGTTTGTACCTTTTGAATCAATTAAATATAGCTTGCTATCAATATATCATCAATTA  | 973  |
| QY | 421  | TTCAAGCGATCGGTAAGCTTTGAAAGTCTTAAGGGTATCAAGCATGCGATCATCGG       | 480  |
| Db | 972  | TTCAAGCGATCGGTAAGCTTTGAAAGTCTTAAGGGTATCAAGCATGCGATCATCGG       | 913  |
| QY | 481  | AAAGATGGCAGCTGTGCTTAAGGGAAGTGCAGATTAAGGCATGTGATAGTATATGCTTGGC  | 540  |
| Db | 912  | AAAGATGGCAGCTGTGCTTAAGGGAAGTGCAGATTAAGGCATGTGATAGTATATGCTTGGC  | 853  |
| QY | 541  | TTTGATTTTTTAATTTGATGATTTGGTAAAGTTTGTTTTTTGTATGATATATATGCAATAC  | 600  |
| Db | 852  | TTTGATTTTTTAATTTGATGATTTGGTAAAGTTTGTTTTTTGTATGATATATATGCAATAC  | 793  |
| QY | 601  | ATCACCATCATCAAAATAATAATAAATCTGCCGTGTGTGCTAAGTAAGCACTATTTGTTGT  | 660  |
| Db | 792  | ATCACCATCATCAAAATAATAATAAATCTGCCGTGTGTGCTAAGTAAGCACTATTTGTTGT  | 733  |
| QY | 661  | GATATATGTGTGATTTTGGGGCAAGGCTGTGCAGTGTGTCAAGATGTGCTTGGCTTAATT   | 720  |
| Db | 732  | GATATATGTGTGATTTTGGGGCAAGGCTGTGCAGTGTGTCAAGATGTGCTTGGCTTAATT   | 673  |
| QY | 721  | GGCATGTCTTGTGTAGATGATGCGGTTTGGCAATTTTCAAAAATAAGCAATTTTGTGTC    | 780  |
| Db | 672  | GGCATGTCTTGTGTAGATGATGCGGTTTGGCAATTTTCAAAAATAAGCAATTTTGTGTC    | 613  |
| QY | 781  | CAAAATAATCCACCATGCGATGTGTGCCATCAAGATGATCAGCGGAGATTTTAAATTTGT   | 840  |
| Db | 612  | CAAAATAATCCACCATGCGATGTGTGCCATCAAGATGATCAGCGGAGATTTTAAATTTGT   | 553  |
| QY | 841  | TGCCCTTGTAGCTGCGCAAAATTGTGATGTGTCAAGGTGAAGAACTTACAGCTCAAGCAC   | 900  |
| Db | 552  | TGCCCTTGTAGCTGCGCAAAATTGTGATGTGTCAAGGTGAAGAACTTACAGCTCAAGCAC   | 493  |
| QY | 901  | CACCAAAATCCATATTTTGAATCGTTAATAATTCAGTGCAGGCGTGCAAATATATCCGCC   | 960  |
| Db | 492  | CACCAAAATCCATATTTTGAATCGTTAATAATTCAGTGCAGGCGTGCAAATATATCCGCC   | 433  |
| QY | 961  | GACACCAACATATATGCTGTATGTTTTGGCATTCGCGCTATCTAATATGTGTGAACAGTGT  | 1020 |
| Db | 432  | GACACCAACATATATGCTGTATGTTTTGGCATTCGCGCTATCTAATATGTGTGAACAGTGT  | 373  |

QY 1021 TTGGGCAATTGAGCTGTGATGCGATGATGCTGTGTTGCTTGAAGTATGCTCG 1080  
DB 372 TTGGGCAATTGAGCTGTGATGCGATGATGCTGTGTTGCTTGAAGTATGCTCG 313  
QY 1081 TGCTTTGAGTGTATCATATGAACTTGAATGATGCTGTGTTGCTTGAAGTATGCTCG 1140  
DB 312 TGCTTTGAGTGTATCATATGAACTTGAATGATGCTGTGTTGCTTGAAGTATGCTCG 253  
QY 1141 GGGGCGCAATATCTCGGCTTCTTGGGTTAATGCCGGGCTGATGATGATGCTCG 1200  
DB 252 GGGGCGCAATATCTCGGCTTCTTGGGTTAATGCCGGGCTGATGATGATGCTCG 193  
QY 1201 TGAGCTTAAGAGTTCATCTGTCCAGCTGCCAAATGCAATGAACGCTGTGGGCACTTT 1260  
DB 192 TGAGCTTAAGAGTTCATCTGTCCAGCTGCCAAATGCAATGAACGCTGTGGGCACTTT 133  
QY 1261 ATGTGCAAGGTAGATTTGGGATTAACCATGATGATGCTTAAGTGTGATGCTGATGAC 1320  
DB 132 ATGTGCAAGGTAGATTTGGGATTAACCATGATGATGCTTAAGTGTGATGCTGATGAC 73  
QY 1321 CAAAAATTGCAAGCTGACAGCTGACAGCTGCCAAACGACGACGATATTTTGTGA 1380  
DB 72 CAAAAATTGCAAGCTGACAGCTGACAGCTGCCAAACGACGACGATATTTTGTGA 13

RESULT 4  
CP000082\_25  
WPCOMMENT

Sequence split into 27 fragments. LOCUS CP000082 Accession CP000082

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| CP000082_00   | 1       | 110000  |
| CP000082_01   | 100001  | 210000  |
| CP000082_02   | 200001  | 310000  |
| CP000082_03   | 300001  | 410000  |
| CP000082_04   | 400001  | 510000  |
| CP000082_05   | 500001  | 610000  |
| CP000082_06   | 600001  | 710000  |
| CP000082_07   | 700001  | 810000  |
| CP000082_08   | 800001  | 910000  |
| CP000082_09   | 900001  | 1010000 |
| CP000082_10   | 1000001 | 1110000 |
| CP000082_11   | 1100001 | 1210000 |
| CP000082_12   | 1200001 | 1310000 |
| CP000082_13   | 1300001 | 1410000 |
| CP000082_14   | 1400001 | 1510000 |
| CP000082_15   | 1500001 | 1610000 |
| CP000082_16   | 1600001 | 1710000 |
| CP000082_17   | 1700001 | 1810000 |
| CP000082_18   | 1800001 | 1910000 |
| CP000082_19   | 1900001 | 2010000 |
| CP000082_20   | 2000001 | 2110000 |
| CP000082_21   | 2100001 | 2210000 |
| CP000082_22   | 2200001 | 2310000 |
| CP000082_23   | 2300001 | 2410000 |
| CP000082_24   | 2400001 | 2510000 |
| CP000082_25   | 2500001 | 2610000 |
| CP000082_26   | 2600001 | 2650701 |

Continuation (26 of 27) of CP000082 from base 2500001 (CP000082 Psychrobacter arcticum 2

Query Match 19.3%; Score 266.8; DB 1; Length 110000;  
Best Local Similarity 53.5%; Pred. No. 7.1e-54;  
Matches 745; Conservative 0; Mismatches 572; Indels 75; Gaps 6;

QY 8 TGAACCATGTTAACTTATCACTGCTCATTTATTTTAAATTTGATCATGCTA 67  
DB 8845 TTAACAGAGTGCTTAAAGGCTGCGCCGCGCAAAACCATTTAACTGATGAGCTT 8904  
QY 68 GCAACAAGAGTGACAAAGACAGCCCTAGCGGTGTTGATGCTGCTGCAATGCTGC 127  
DB 8905 GCACAAGAGTGACATATATACCGCTGTACTTGCGAACAACTACTGCGCTTACTTGC 8964  
QY 128 CGA-----TCAAGTACCAACCTGATATATTTTAAATGCCAAATGT 169

DB 8965 TCAATCATGAAAAAGCAATTTTCTAAGCTGTGATGATGACGCTAACTTCATGCTG 9024  
QY 170 GCATGACCTTTAGTAAATCTTGCTCAATTAATCGGGCATCTTGAACGATTAATACACA 229  
DB 9025 AGACCCGCTGCGGTAAATGCTGTTCAATCTGTGCTGCTCTCTCGCAATTAAGACACT 9084  
QY 230 CTATGACCACTGATTAACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAG 289  
DB 9085 TGACTCACATATGATTTGAATTAAGGCGAGAGTTGACCGAAGCTGAGCTTACCTTGA 9144  
QY 290 CCACCCAAATCAATGCCA-----GTATGATATGCCATTAACCTGCGCAACCATCA 343  
DB 9145 CCGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9204  
QY 344 ATGTGCAATGCTTGAAGCGATGTTTGTACCTTTGATCAATTAATTAATTAATTAATTAAT 403  
DB 9205 ACAGCCGCAATGTTTGAAGCAATATTTGCTTTTGAATTAATTAATTAATTAATTAAT 9264  
QY 404 TCAATATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 463  
DB 9265 GCAATTTGGGCGCATATTTGAGAGCGATGCTCAAGACTGCAAACTGTGCAAGGTGCTA 9324  
QY 464 AGCATGCAATTCATGAGGAGATTTGGAGCTGTGCTTAAGCAAGTGCATTAAGGATG 523  
DB 9325 AGCATGCTCTCAAGCGGCAAAACGACAGCTGCGCCCAATGCAATGCTGCTGGGCTATTA 9384  
QY 524 AGTAGGTATGCTTGTGCTTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 583  
DB 9385 AGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9444  
QY 584 TGATATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 643  
DB 9445 CGTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9485  
QY 644 AGTAACCTATTTGCTGCTGATTAATGCTGATTTGGGCAACGCTTGTGCTGCTGCTGCTGCT 703  
DB 9486 ATAGTCCCTTGTGATGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9545  
QY 704 CATGCTTGCTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763  
DB 9546 CATGCTTGCTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9603  
QY 764 ATAGCAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823  
DB 9604 ATAGCAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9663  
QY 824 GAGAGTTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
DB 9664 GAGAGTTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9723  
QY 884 CTGTAAGCTCAAGCAACCAATTCATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943  
DB 9724 CTGTAAGCTCAAGCAACCAATTCATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9783  
QY 944 GTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1003  
DB 9784 AGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9843  
QY 1004 AATGCTGTAAGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063  
DB 9844 AATGCTGTAAGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9903  
QY 1064 GCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1123  
DB 9904 GCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9936  
QY 1124 GGAATGCTTGAAGCTTTGGCGGCGCAATATCTGCGGCTTTGCTGCTTAAATGCGGGGCTG 1183  
DB 9937 GGAATGCTTGAAGCTTTGGCGGCGCAATATCTGCGGCTTTGCTGCTTAAATGCGGGGCTG 9996  
QY 1184 ATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1240  
DB 9997 ATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10056





Db 70234 TGAATGCGCCAGTGAAGTCTCTCGAATCTGTGACAGCAACGCGGTGCGCTTGTTCA 70175  
QY 1319 ACCAAAAATTTCAGCTGACAGTGCAGAGCTGCCCAACGAGAGCGAT 1370  
Db 70174 TGTAAAAATTAAACGGCTGAACCCCTTGAGATTCAGAGACCCGCAAGCACTT 70123

RESULT 6  
CP000082\_24  
WPCOMMENT

Sequence split into 27 fragments LOCUS CP000082 Accession CP000082

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| CP000082_00   | 1       | 110000  |
| CP000082_01   | 100001  | 210000  |
| CP000082_02   | 200001  | 310000  |
| CP000082_03   | 300001  | 410000  |
| CP000082_04   | 400001  | 510000  |
| CP000082_05   | 500001  | 610000  |
| CP000082_06   | 600001  | 710000  |
| CP000082_07   | 700001  | 810000  |
| CP000082_08   | 800001  | 910000  |
| CP000082_09   | 900001  | 1010000 |
| CP000082_10   | 1000001 | 1110000 |
| CP000082_11   | 1100001 | 1210000 |
| CP000082_12   | 1200001 | 1310000 |
| CP000082_13   | 1300001 | 1410000 |
| CP000082_14   | 1400001 | 1510000 |
| CP000082_15   | 1500001 | 1610000 |
| CP000082_16   | 1600001 | 1710000 |
| CP000082_17   | 1700001 | 1810000 |
| CP000082_18   | 1800001 | 1910000 |
| CP000082_19   | 1900001 | 2010000 |
| CP000082_20   | 2000001 | 2110000 |
| CP000082_21   | 2100001 | 2210000 |
| CP000082_22   | 2200001 | 2310000 |
| CP000082_23   | 2300001 | 2410000 |
| CP000082_24   | 2400001 | 2510000 |
| CP000082_25   | 2500001 | 2610000 |
| CP000082_26   | 2600001 | 2650701 |

Continuation (25 of 27) of CP000082 from base 2400001 (CP000082 Psychrobacter arcticum 2)

Query Match 16.6%; Score 229.6; DB 1; Length 110000;  
Best Local Similarity 54.3%; Pred. No. 7.2e-45;  
Matches 581; Conservative 0; Mismatches 444; Indels 45; Gaps 4;

QY 8 TGAACCATGTTAAACAATTATCACTCGCTCATTTAATTTTAAATGATCATGCTA 67  
Db 108845 TTAAACGAGCTGCTTAAGCGCTGCCCCGCGCAACAACCATTAACATGATGAAGCTT 108904  
QY 68 GCACAGCAGGTGACAAAGCAGCCTAGCGGTCTTGACTGGCTGCCAATACCTGCC 127  
Db 108905 GGCACAGCAGGTGACAAATACCGCTGACTTGGACAACTACTTGGCTTACTTGC 108964  
QY 128 CGA-----TCAGTGTACCAACTGATATATTTTAAATGTCGAATGT 169  
Db 108965 TCATCATGAAAAACATTTTCTAAGTCTGACATTTGACAGGCTAACCTCATGCTG 109024  
QY 170 GCATCACTTTGTAATCTTGTCAATTTATGAGGAGCATCTTGACGATTAATACACA 229  
Db 109025 AAGACCGCTGCGGTAAAGTCTTTCAATCTGCTACTGTCTCGCAATAAAGAGCACT 109084  
QY 230 CTATCCACCACCTGATTAACATATGCTGACAGCTCACTAAATCTTGACCTTACCAAC 289  
Db 109085 TGACTCACAATATGATTAAGAAAGGAGAGTTTACCGAAGCTGAGCTTTACTTTGA 109144  
QY 290 CCACCCAAATCAATGCCA-----GTATGATATGCCATAAACCGTCCAGGCCATCA 343  
Db 109145 CCGCCCAATTAATTAATATTAATTAACCTGTTGGGTGCAATAGACCCGCCCAACCTTCA 109204  
QY 344 ATTGTGCAATGGTGTAGCGCATGTTGTATCCTTTGAATCATTAATAATATGCTTGTA 403  
Db 109205 AAGCGCCGATGGTGTAGCCAAATATGTGCTTTAGAAATCATTAAGTAATCAATGCCA 109264

QY 404 TCAATATCATATATATTTACAGCGATGCGGTAGCCCTTTGAAAAGTCTTAAGGTATCA 463  
Db 109265 GCAATTTGGCGACATATTTGGCAGAGATGCTCAAGACTGTCGAAACCTGCTGCAAGGTGTA 109324  
QY 464 AGCATGCAATTCATGGAAGATTTGGAGCTGTGCTTAAGGCAAGTGCAGTAAGGCAATGG 523  
Db 109325 AGCATCTCTCAAGCGGCAAAACCGACAGCTGCGCCCAATGCCAATGCTCTGGGATTA 109384  
QY 524 AGTAGGTATGCTGCTTGAATTTTAAATGATGATGATGTAAGTTGTTTTTTTGA 583  
Db 109385 AGTAAGTATGACGACCTTTTATCTTAATTTATCTGTAATGAGACGCTCTGTACCA 109444  
QY 584 TGAATATATGCAATACCATCACCATCATCAATTAATAAATATCCGCTTTGTTGGCTTA 643  
Db 109445 CGTGCCAAATAGTCTGCGCTGCTGATCCGTAAT-----AGCCC 109485  
QY 644 AGTAAGCTATTTGTTGTGTATTAATGTGTGATTTGGGCAACGCTTGTCAGTGTCAAG 703  
Db 109486 ATAGTCCCTTGTGAGAGCATGAAATACAGTGTCACTGCGGCAACATATCTGACAC 109545  
QY 704 CATGCTGGCTAAATTTGGCATCGTCTGTGATGATGATGCGGTTTGGCAATTTTCAAA 763  
Db 109546 CAATGTCGGGTCA--CGATCTTCACATTAATCACTACGATTTGACCTTGAAG 109603  
QY 764 ATACGCAATTTTGTGCAAAATATCAACCATGCTGTCGCGCATCAAGATGATCAGCG 823  
Db 109604 ATACGATTTTGGCTTGTGATGATGAGCGACATATACATGACGCTTAATGATCAGCG 109663  
QY 824 GAGAGATTTAAATTTGTTGCCCTTGAAGCTGCGCAATTTGAGATATGCTCAAGCTGA 883  
Db 109664 GACATATGAGACCTGTTGCGACCTGCGCCCTTAAGTTGTGACGTTTCTAATCTGAAG 109723  
QY 884 CTGACAGCTCAAGCAGCAACAATCATATTTTGAATGTTAATTAATCAAGTGCAGCG 943  
Db 109724 CTAGATATTTCTATTAATCTGCTGCTCATTTTATCTAATAAGCTCAGCGCTGCG 109783  
QY 944 GTGCCAATTTACCGCGCAGACCAACAATCATGCTGATGTTTGGCATCTGCGCTACT 1003  
Db 109784 AGCGCAATTTACCGCTGACGCAACATTTGATGACGATCTGCGGCACTTTGTCTCAAC 109843  
QY 1004 AATGTGTACAGTGTGCTTTGGCATTTGAGCTGTGATGCGCATGATTTGG 1053  
Db 109844 AGCGTGTACAGTGTGCTTTTGCATTTGAGCCAGTAATGCAACAATTTGG 109893

RESULT 7  
ARJ18732/c ARJ18732 1416 bp DNA linear PAT 17-AUG-2003  
LOCUS  
DEFINITION Sequence 1282 from patent US 6562958.  
ACCESSION ARJ18732  
VERSION ARJ18732.1 GI:33699835  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1416)  
Breton, G. and Bush, D.  
Nucleic acid and amino acid sequences relating to Acinetobacter  
baumannii for diagnostics and therapeutics  
Patent: US 6562958-A 1282 13-MAY-2003;  
JOURNAL  
Genome Therapeutics Corporation; Walham, MA  
FEATURES  
source  
1..1416  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 16.4%; Score 226.2; DB 6; Length 1416;  
Best Local Similarity 52.1%; Pred. No. 8.6e-44;  
Matches 718; Conservative 0; Mismatches 593; Indels 66; Gaps 7;  
QY 2 AAAGCGTGAACATGTTAAACAATTATCACTCGCTCATTTAATTTTAAATGATGATCC 61

```

Db      1409 AACCAATTGACGACGAGCAACAACTGCTGACACGCGTCAATTAATCTTTTAAACATATCA 1350
Qy      62   ATGCTAGACAAAGAGGAGCAAAAAGCAGCCCTAGCGGCTTTGACTGGGCCCAAT 121
Db      1349 AAACCTTGACACATGCTGTGAT-----AGCAATACCAATCTTACGCTGTGTTTAC 1298
Qy      122   ACTGCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAATGTGATCACTTGG 181
Db      1297 GTTGACACAGCTGACAGCTTCTTTAGCGTTGCTGATGTAATAATTTAGTTCGGCT- 1239
Qy      182   AGTAATCTTGCTCAATTAATCGGGCAATTTGACCGATTAATAACACATATCAACCCAC 241
Db      1238 TGAATGCTGTGCTGATGACAGCGGCAATTTCAACAAATACCAACATTTGGCATAT 1179
Qy      242   TGCATTAATATGCTGACAGCTCACTAAATCTTGACTTGAACCAACCCCAATATC 301
Db      1178 TTTTCATAGAAAGACCTTAAGAGAAATAATCTTGTCCCTTGGCCCTGTCCATTAATA 1119
Qy      302   AATGCCA-----GTGATGATATGCCATTAACCGTCCAGCCCATCAATTTGCTGAATG 355
Db      1118 AGTGCAACCTTACTCTTTTCACTTCAATGACGACACCTTAAGCATATTTGCTGCAAGT 1059
Qy      356   GTTGAGCGGATGTTTGTACTCTTTGAATCATTAATAATATGCTTGTATCATATATCA 415
Db      1058 GTAGCAACCAATAGTACCTTTAGAGTCAATTAATAAGCAATCATGACGCTTTTA 999
Qy      416   ATATATTCACAGGAGCGGAGAGCCCTTTGAAGTCTTAAGGATTAACATGACGATCC 475
Db      998   ACGTACTCACAGGCGTCTCTAGCTCTTAAATTTTAAAGTTTAAACATTTAGCTCC 939
Qy      476   ATGGAAGATTGGCAGCTGTGCTTAAGGCAAGTGACAGATAGGACATGAGTATGATG 535
Db      938   ATAGTAAACCAATGTGCTTCCCAATGCTTAACACTTAAGATTAAGTACATTTGCTGC 879
Qy      536   TTGCTTTGATTTTAAATGATGATGTTGTAAGGTTTGTGTTTGTGATATATATGCC 595
Db      878   ATACCTTGATATATTAATCTGAGCTTTTAAATTAACCGCTGTAACACAGACAGCA 819
Qy      596   ATACCAATCAACATCAATTAATAATAATAATCTGCCCTTTGCTGCTTAAGTATATTT 655
Db      818   AGCGTACCGTGGCATCTTTAAAGCGCATATTTGAT-----TAAATCCGCTGCGT 767
Qy      656   GTTGTATATATGTTGATTTGGGGCAACGCTTGTCAAGTGTGACAGATGCTTGCGCT 715
Db      766   TTAAACCAAGCTTTGATTTGTGTCTATCTGAAACAATGTGACGCTTA----- 716
Qy      716   AAATTTGCAATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAAATACGCAATTT 775
Db      715   ----ACGATCATCTCGGTTAAATAACAATTTTACGCGCTTGAAAAATACGATGTTT 660
Qy      776   TGTGCCAAATATTCACCAATGCCATGCTGCCATCAAGATGATCAGGAGAGATTTAA 835
Db      659   GCTTATGATATACCCAGCATATTTCAATGACGCTTAATGCTTCACTCATATTTGAGA 600
Qy      836   ATTTGTCCTTGAAGCTGACCAATTTGAGATATGCTCAAGCTGAAAACTTGAAGTCA 895
Db      599   ACACCTGTACTCTCAGGCTTTAAGTGAAGTGTGTTTCTTACTGAAAGCTTGTAATCTA 540
Qy      896   AGCAACAACCAATTCATATTTGATGCTTAATTAATTAATGATGCAAGCGCTGCCAATAT 955
Db      539   AGGACAGATACCTG---TTGATCTTTAAGTAATCTTAACGCTGGTCAACCAAGTTG 483
Qy      956   CGGCGACACCAACATCAATGCTGCAATGTTTGGCATCTGCGCTACTAATGTCTGAACA 1015
Db      482   CGGCTACTAGCAATTTTCTTACTGATCTTGAAGCAATTAAGCCAAATTAAGGTTTACA 423
Qy      1016   GTGCTTTTGGCATTTGAGCTGTGATGAGCGATGATGTTGTGTTGTGCTTGAAGTTAGA 1075
Db      422   GTACTTTTGTGATTTGAACTGTATATGCAAAATGCGACATCAAGTGC----- 373
Qy      1076   TCTGTGCTTTGATGTATCAATGAACAATTTGAACATGCTGTATCAAGGAATGCTTGA 1135

```

```

Db      372   -----ACGCGCAATTAATCTGATATGCCCCACACAGAAATACCTTTA 330
Qy      1136   GCTTTGGCGGACCAATATCTGGGCTTCTTGGGTTATGCGGGGCTGATGATGATCTGA 1195
Db      329   GCAATAGCTCTCGAATTTCCGTAATTTGTGTGCAAGGCGCTGGCTTAAATAATCTCT 270
Qy      1196   TCGGCTGAGCTTTAAGATTCTCTGCTCCAGCTCCAAATATGACAAATGACGCGTGGGC 1255
Db      269   TCTGCTGTATATTAATTTCTTGTATCAAGCTGACCAAACTGCTTTTAACACCGCGGA 210
Qy      1256   AGTTATCTCCAAAGTAGGATTTGAGATTACATGATGATGCTTAATTTGTGACCGTGA 1315
Db      209   ATCTGATCGTGTCCGGAGGTGTGGGCGGGAATCCGTTACAGCAACTTGTAGCTTGT 150
Qy      1316   TTGACCAAAATTTTCAAGCTGACAGTCCAGAGCTGCGCCAAACGACAGAGGATAT 1372
Db      149   TCATGACAGAAATTTACAGCAAAACCTGATATTTCCAAAGCTGTCTACAACTTTT 93

RESULT 8
AX932204/c 1314 bp DNA linear PAT 22-DEC-2003
LOCUS      Sequence 157 from Patent W003087353.
DEFINITION AX932204
ACCESSION  AX932204
VERSION     AX932204.1 GI:40312625
KEYWORDS
SOURCE
ORGANISM   Haemophilus influenzae
            Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE
1 Edwards, A., Dharamsi, A., Vedadi, M., Domagala, M., Houston, S.,
  Avery, D., Beatrice, B., Mansoury, K., Ouyang, H., Vallee, F.,
  Richards, D., Neherly, K., Virag, C., Banadziya, K., Pinder, B.,
  Alam, M.Z., Tai, M., Camadlen, V. and Kanagarajah, D.
  Patent: WO 03087353-A 157 23-OCT-2003;
  Affinium Pharmaceuticals, Inc. (CA)
FEATURES
source
1..1314
   /organism="Haemophilus influenzae"
   /mol_type="unassigned DNA"
   /db_xref="taxon:727"

ORIGIN
Query Match      11.6%; Score 160.4; DB 6; Length 1314;
Best Local Similarity 49.6%; Pred. No. 7,5e-28;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy      1   TAAAGGTGAACCATGTTAAACAACCTTACCTGCTCATTAATTTTAATTGATC 60
Db      1308   TAAACATTGAGCTAAATGCGTAAATCTTTCGCGCTTTTCAAAAGAAACCAATGATC 1249
Qy      61   CATGTAGCAACAAGCAGTGACAAAAGCAGGCTTAGGCGTCTTGACTGGCTGCCAA 120
Db      1248   GAGACTTGACAAAGAGGCGAGCAATATACA-----TATCTCGCTTGCA 1201
Qy      121   TACTGCCCGATCAAGTGTACCAACTGATATATTTAGATGCCAAATGTGATCACTTT 180
Db      1200   TGTGGGCGTAAAAATTTTATGCTGTTCATTTGATGTAACAA-----ATTAATTTG 1147
Qy      181   GAGTAATCTTGTGCTCAATTAATGCGGAGCATTTGACCGATTAATAACACTATCAACCA 240
Db      1146   CGATGAAATTTTGGACAGCGGCGACCACTTGACCAAAACATTAATTAATGTGTGG 1087
Qy      241   CTGCAATACATATGCTGACAGCTCACTAAATCTTTGACCTTTACCAAGCCCAACCAAT 300
Db      1086   TTGATTAATTAATTCAGTAAATTTTGAAAAATACAGCCCTTTTCCGCTCCGCTAGCA 1027
Qy      301   CAATGCAATGATGATATATCCATTAACGCTGCAAGCCCATCAATTTGTGCAATGTTGA 360
Db      1026   CAATGCAATTAATCCCTCAATATA-----AAGCCAGGCAATGACAGCAACTGTACT 976
Qy      361   GCCGATGTTTGTACTTTTGAATCATTAATAATATGCTTGTGATCAATATCATATATA 420

```

Db 975 CCCACATTTGTTGCTTAAAGTCATTAATCCAAAGCATCCATTAGCTGATGCACTAA 916  
421 TTCAAGCGATGGGATGAGCCCTTTGAAAGCTTAAGGGATTCAGACATGCAATCGG 480  
915 TTCAAAACATGATCTTAACCTTTGAAATACGAGGCGGTCGAAATTAATCTAATT 856  
481 AAGATTGGCAGCTGTGCTTAAGCAAGTGCAGATTAAGCATTAAGTATGCTTCC 540  
855 AATACCTAATAGCTGTGCGCAATGCTGTGCGCAAAATGTTCAATTAATTAAGGAGC 796  
541 TTGATTTTAAATTTGATGATTTGATTAAGTTGTTTGTGATGATTAATGCAATCC 600  
795 AACCAATGATGCTTCTTCAACAGATAAATCACTTCACTTTTAACATTAATTAATGCTT 736  
601 ATCAACATCATCAATTAATTAATTAATCTGC-CCGTTGCTGCTTAAGTATGCTTGTG 659  
735 GCATTTTCACTTTTAAAGCAATTAATCCGCACTATTTCCGCAAAAGAAACGTAATGTT 676  
660 TGATTAATGATGATTTGGGGCAACGCTTGTCAATGATGATCAAGCATTTGCTGCTAAT 719  
675 CGCTTATTTTGTGTTTTCCCAAAAGTCAGCC----- 644  
720 TGGCATGCTTGTGATGATGATGCGGCTTGGCAATTTTCAAAATTAAGCAATTTTGTG 779  
643 ---TATCTTCAATGTTTCAACACACTTACCTTATGATTAATGCTTAATTTTGTCTT 587  
780 CCAATTAATCACCATGATGCGCATGCGCATCAAGATGATGAGGAGATTTAAATTTG 839  
586 GGGGATTAATCTTTAAATCATTAATGAGGATCAATGATTTTCAATCACTTCAAGACG 527  
840 TTGCCCCCTTGAAGTGCCTTGAATTTGATGATGCTCAAGCTGAAATTAAGCTCAAGCA 899  
526 TGGCTGCGGAGCTTTTAAAGCTTAATGATGCTTCAAGCTGAAATTAAGCTCAAGCA 467  
900 CCACCAATTCATTTTGTGATGCTTAATTAATCAAGTGAAGCGCTCAATTTTACCG 959  
466 CATTAAGTTCATTTTCAATTCATTC---ACATTAACAAAGGGAATCCCAATTTTCCG 410  
960 CGACACCAACATCATGCTGCTGATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019  
409 CCATACCAATTTTCAACACGAGCTTTGCGCAATTTCAATTAATTAATTAATTAATTA 350  
1020 TTTTGGCAATTTGAGCTGTGATGCGATGATGATGATGATGATGATGATGATGATG 1079  
349 TTTTACATTTTGAACCTGTAATCCCAATTTGCTTGTGCTGCTGCTGCTGCTGCTG 298  
1080 GTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139  
297 -----GAAATTAATCAATATGCGGATTAATTTTCACTCCGCTTTTAA 257  
1140 TGGCGGCAATTAATCTGCGGCTTCTTGGGTTAAATGCGGCGGCTGATGATGATGATG 1199  
256 GTGCGGTTGAATTTTGTGCTTAAATGCGGAGGCTTGGGCTTAATTAATTAATTAAT 197  
1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259  
196 TTTCAATTAATCAATCTGCTGATTTAAATCAATGATTAAGAGGATTAATTTTGAAGA 137  
1260 TATCTGCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319  
136 TATCAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77  
1320 CCAAAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
76 AAGATTAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 19

RESULT 9  
U32793 11498 bp DNA linear BCT 02-JUN-2004  
LOCUS U32793/c  
DEFINITION Haemophilus influenzae Rd km20 section 108 of 163 of the complete genome.

ACCESSION U32793 L42023  
VERSION U32793.1 GI:1574683  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Merrick, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shiley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Sauder, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrman, J.L., Geophagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.  
Whole-genome random sequencing and assembly of Haemophilus influenzae Rd  
Science 269 (5223), 496-512 (1995)  
JOURNAL  
PUBMED  
7542800  
REFERENCE  
AUTHORS  
Tatusov, R.L., Muehlegan, A.R., Bork, P., Brown, N.P., Hayes, W.S., Borodovsky, M., Rudd, K.E. and Koonin, E.V.  
Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli  
Curr. Biol. 6 (3), 279-291 (1996)  
JOURNAL  
PUBMED  
8805245  
REFERENCE  
AUTHORS  
White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.  
Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA  
4 (bases 1 to 11498)  
JOURNAL  
PUBMED  
8805245  
REFERENCE  
AUTHORS  
White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.  
Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA  
4 (bases 1 to 11498)  
JOURNAL  
PUBMED  
8805245  
REFERENCE  
AUTHORS  
White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.  
Direct Submission  
Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA  
The whole genome was shifted by 588 nucleotides for a new start on Oct 1, 1996 this sequence version replaced gi:1221887.  
LOCATION/Qualifiers  
1..11498  
/organism="Haemophilus influenzae Rd km20"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:71421"  
131..586  
/gene="HI1129"  
131..586  
/gene="HI1129"  
/note="similar to SP:p2186 GB:X52063 PID:40849 PID:42319 GB:U00096 percent identity: 59.21; identified by sequence similarity, putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC22784.1"  
/db\_xref="GI:1574684"  
/translation="MPGATAVNLDSKRVATPPRYAETLEKNGOMVCTVDIPDSC LIIYFDDEKTEKGLALNSNDPIORLORVMGHAETCEKDAQRIILSGPLRQHA KIEKLMVVGQNLKFEIWSDVEMHQIADIEIGSTDFADALNDFSL"  
617..1582  
gene

gene  
CDS  
617..1582  
/gene="H11130"  
/note="similar to GB:D10483 SP:P18595 GB:X52063 PID:216497  
PID:40850 percent identity: 62.09, identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAC22785.1"  
/db\_xref="GI:1574685"  
/translation="MNSSESPSSSHITVLLHEAVNGSLKENGIDYIDGTFGRGSHR  
RLISDLSNGRIADVDRPRAIRABAKIKQDARFQENHSFHHIPICKLANVKTIOG  
IILDGVSPFDLEABRGFSFMKQDPLMRMDITDGLSABEMLKQVSLIEDITVLLKTFP  
GGERAKRATAVDPFNKSAVNKGTEFISRSQALIELSQAVPFRDKHPRATSFOA  
IRIFNSELDELSLALDMLAPEGLSLISFSLDPRVAKFMKKQSGEIDPKG  
LPLRDEIQORNOCKRTIKGAIQPDAAEQIAMPRSRAILRAERT"  
1585..1908  
/gene="H11131"  
1585..1908  
/gene="H11131"  
/note="similar to SP:P22187 GB:K00137 GB:S49802 GB:S49875  
GB:X52063 percent identity: 36.63, identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="cell division protein (ftsI)"  
/protein\_id="AAC22786.1"  
/db\_xref="GI:1574686"  
/translation="MSENNKRPRYPQQLVLEDLBSNGLVYLLIGLVSAMGITWTT  
HKTRQLISENGLILQROLKENRYNNLQVQETBGDSITRESIAISTLKRVSSDEGE  
VEIR"  
1921..3753  
/gene="H11132"  
1921..3753  
/gene="H11132"  
/note="similar to GB:D10483 SP:P04286 GB:K00137 GB:S49802  
GB:S49875 percent identity: 53.11, identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="penicillin-binding protein 3 (ftsI)"  
/protein\_id="AAC22787.1"  
/db\_xref="GI:1574687"  
/translation="MVFNSSRKSGKSKKTRKLTAPETVKNKPKVFEKCFMRGRY  
MLSTYLLIGLCALVABAAYOSINADLTLSNADKRSIRKQEVLSVSGSITDRGOLL  
SVSVMSMIVADRPKMLKENSADKERIALAAEELGMENTDLYKTIKNSGSLYTLA  
ROYELSKANYRRLKIKGIIILETHRRRTPEVEEAHVAGTDDIGNGIESEKSFNS  
LIVGQSGRTVARKDGRGNIVAHISDEKTYDQDVLTLSDEKLGSMYAREIKASENN  
AASGAVLAVDVATGEVLAMATAPSPNNRIVGVSGSLAKRNALIDTFEPGTVAPFVAV  
LTALORGVKDEIDITTSFKLSGEIVDVAQAOULDEILMNSNRGVSRAPIARMP  
PSALMETYQNGLSKRPDGLIGBOVGIILANRRKMDIBERTVAVGVTATPRLQIA  
RAYALTSFGVRPLSTTKDPPVYGRVSEKTKTQVGLIEKVALIKRRKAMVEGR  
VGVTGRKRIKENGHTVAKTVAFTAGLIPISDPYALVLIINDPABEYIGGAVSAPV  
FSNINGLVALRNAVLPQDAEAEMTTTSKAKIYVIGEHQKN"  
3763..5229  
/gene="H11133"  
3763..5229  
/gene="H11133"  
/note="similar to GB:D10483 SP:P22187 GB:X55814 PID:285768  
PID:581032 percent identity: 55.72, identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="UDP-N-acetyl-muramyl-tripeptide synthetase  
(murB)"  
/protein\_id="AAC22788.1"  
/db\_xref="GI:1574688"  
/translation="MKKLTALFNLPELKNIDIELHNVULDSKAGDLFYAIKSHQVD  
GNQFDLSALHSGASAVVSEETLSSEHLTVARIGVPPVYKQLAAHLSLADVYDVS  
SNMLTVGYTGNGKTTISQLAQAEELLGRBAVWGVTIGNGLFGQIVKANTQVSSAV  
EIOSSLSPKHAQADFTSIEVSSHGLAQRVEALHPKALIFTNLTRDLIDHOSMEY  
AAAKRLLTELDTOIKVINADDEIGYQWLTLPDAIVASNNADPKVSGHOMKQINHY

| Query Match | Best Local Similarity | Matches | 683;   | Conservative | 0; | Mismatches | 606;   | Indels | 89; | Gaps | 7; |
|-------------|-----------------------|---------|--------|--------------|----|------------|--------|--------|-----|------|----|
| 1           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 2           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 3           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 4           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 5           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 6           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 7           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 8           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 9           | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 10          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 11          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 12          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 13          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 14          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 15          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 16          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 17          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 18          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 19          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 20          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 21          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 22          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 23          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 24          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 25          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 26          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 27          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 28          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 29          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 30          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 31          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 32          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 33          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 34          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 35          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 36          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 37          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 38          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 39          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 40          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 41          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 42          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 43          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 44          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 45          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 46          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 47          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 48          | 11.6%;                | Score   | 160.4; | DB           | 1; | Length     | 11498; |        |     |      |    |
| 49          | 11.                   |         |        |              |    |            |        |        |     |      |    |

Db 8900 TTGATTAAATTAATTCAGCTTAATCTGAAAAATCAGCCCTTTCCGCTCCGCTGACAA 8841  
QY 301 CAATGCCAGTATGATATGCTATAAACCGTGCAGGCCATCAATGCTGCATGTTGA 360  
Db 8840 CAATGCAATTTACCTCATATA-----AAGCCCAAGCAATGCAACTGTA 8790  
QY 361 GCCGATGTTTGAATCTTTGAAATCAATTAATATGCTGCTATCAATATCATCAATATA 420  
Db 8789 CCCCAATTTGTTGCTTTAGATCAATTAATCAAGCAATGCTATGCTGATGACATA 8730  
QY 421 TTACAGCGATGCGGTAGCCCTTTGAAAGCTTAAGGATCAAGCATGCGATCCATGG 480  
Db 8729 TTGAAAAAGATGATCTAAACCTTTGAAATACCAAGGCGGTGCAATTAATCTAAAT 8670  
QY 481 AAGATTGGCAGCTGCTGAGGCAAGTGACATTAAGCATAGATGATGCTTGGC 540  
Db 8669 AATACCTATAGCTTTGCGCAATGCTGTGCTGCAAAATGTTCAATTAATATGCGAC 8610  
QY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 8609 AACCAATGATGCTTTTCAAGATTAATCACTTCACTTTTACATTAATATGCTT 8550  
QY 601 ATACACATCATCAATTAATTAATTAATCTGC-CGTTGGTGGCTAAGTAACTATTTGGT 659  
Db 8549 GCCATTTTCAATTTTAAAGCAATTAATCCGCAATATTTTCCGCAAAAGAAACGGTATGTT 8490  
QY 660 TGAATAATGATGATTTGGGGCAAGCTTTGCAATGATGATGATGATGATGATGATGAT 719  
Db 8489 CGTTGATTTTCTTTTCCCAAAAGTCAAC----- 8458  
QY 720 TGGCATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779  
Db 8457 ---TATCTCATTTGTTCAACACACTACTTATGATGATGATGATGATGATGATGATGAT 8401  
QY 780 CCAATTAATCCACATGCGATGCGGATCAAGATGATGATGATGATGATGATGATGATGAT 839  
Db 8400 GCGAATTAATCTTAAATCAATATAGCATGATGATGATGATGATGATGATGATGATGAT 8341  
QY 840 TTGCCCCCTGAGCTGCAATTTGATATGATATGCTCAAGTGAATTTGACAGCTCAAGCA 899  
Db 8340 TGGCTGCGGAGCTTTTAAAGCTTAATGATGATGATGATGATGATGATGATGATGATGAT 8281  
QY 900 CCAACCAATTCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959  
Db 8280 CATTAAGTTCAATCTTCAATC--AACAATGACAAAGGGGAATCCCAATTTTCCG 8224  
QY 960 CCAACCAATTCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
Db 8223 CCAATCAACATTTTCAACAGCAGCTTTGCGCATTTTCAATTAATGATGATGATGATGAT 8164  
QY 1020 TTTTGGCATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
Db 8163 TTTTACCAATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8112  
QY 1080 GTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139  
Db 8111 -----GAATTAATTCATATTCGCGATTAATCTTCACTCCGCTTTTA 8071  
QY 1140 TGGCGGCAATATCTCGGCTTTCTTGGGTTAATGCGGGGCTGATGATGATGATGATGAT 1199  
Db 8070 GTGCGGTTTGAATTTCTGATGATTTTAAACGCAAGCCCTGGGCTTAATTAACATCATAT 8011  
QY 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
Db 8010 TTTTCAATTAATCTTCTGATTTTAACTACAGTATGATGATGATGATGATGATGATGATGAT 7951  
QY 1260 TATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319  
Db 7950 TATCAATATCAAGATGATGATTTTTCAGATTAATCAATCAAGATTAATGATGATGATGAT 7891  
QY 1320 CCAAAAAATTCACAGCTGACGCTGACAGAGCTGCCAAACGACGACGCAATTTT 1377

Db 7890 AGAGATTAATCCACACAAAGAAAGCTGTTTGGCAAGCCCGATGATGATTAATTTT 7833  
RESULT 10  
AX191762/c 11498 bp DNA linear PAT 15-AUG-2001  
LOCUS AX191762 Sequence 44 from Patent WO0149775.  
DEFINITION AX191762  
ACCESSION AX191762  
VERSION AX191762.1 GI:15209931  
KEYWORDS  
SOURCE Haemophilus influenzae  
ORGANISM Haemophilus influenzae  
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
Pasteurellaceae; Haemophilus.  
REFERENCE  
1 Iversen, P. L.  
Antisense antibacterial cell division composition and method  
JOURNAL Patent: WO 0149775-A 44 12-JUL-2001;  
Avi Biopharma, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..11498  
/organism="Haemophilus influenzae"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:727"  
ORIGIN  
Query Match 11.6%; Score 160.4; DB 6; Length 11498;  
Best Local Similarity 49.6%; Pred. No. 5.6e-28;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
QY 1 TAAAGGTGAACCATGTTAAACAATTAATCACTGCTCATTAATTTTAAATGATC 60  
Db 9122 TAAACATTAAGCTTAATGCGTAATTTCTGCGCGCTTTTCAAAAGAAAGCAATGATC 9063  
QY 61 CATGCTAGACAAGAGGTGACAAAGCAACAGCCCTTAAGGCTGTTGATGCTGCTGCA 120  
Db 9062 GAGACTTGCACAAGCAGGCGCAATTAATCA-----TATCTCCGCTTTGCA 9015  
QY 121 TACTGCGCATCAAGTATCAACCTGATATTTTAAATGCAATGATGATGATGATGATGAT 180  
Db 9014 TGTGGGCTTAAATTTCTATGCTTTTCCATTTGATGACAA-----ATACTTTG 8961  
QY 181 GAGTAATCTTGTCTCAATTAATGCGGATCTTGAACGATTAATTAATCACATATCCACCA 240  
Db 8960 CGATGAATATTTTGCAGAGCAGCAACATCTGACCAAAACAAATTAATATGATGATG 8901  
QY 241 CTGATTAATATGCTGACAGCTCAATTAATCTTGAACCTTTTCAAGCCCAACCAAT 300  
Db 8900 TTGATTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8841  
QY 301 CAATGCCAGTATGATATGCTATAAACCGTGCAGGCCATCAATGCTGCATGTTGA 360  
Db 8840 CAATGCAATTTACCTCATATA-----AAGCCCAAGCAATGCAACTGTA 8790  
QY 361 GCCGATGTTTGAATCTTTGAAATCAATTAATATGCTGCTATCAATATCATCAATATA 420  
Db 8789 CCCCAATTTGTTGCTTTAGATCAATTAATCAAGCAATGCTATGCTGATGACATA 8730  
QY 421 TTACAGCGATGCGGTAGCCCTTTGAAAGCTTTAAGGATTAAGCATGCGATCCATGG 480  
Db 8729 TTGAAAAAGATGATCTAAACCTTTTAAATGACGAAGTGGGATCAATTAATGATTAAT 8670  
QY 481 AAGATTGGCAGCTGCTGAGGCAAGTGACATTAAGCATAGATGATGATGATGATGATGATG 540  
Db 8669 AATACCTATAGCTTTGCGCAATGCTGTGCTGCAAAATGTTCAATTAATATGCGAC 8610  
QY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 8609 AACCAATGATGCTTTTCAAGATTAATCACTTCACTTTTACATTAATATGCTT 8550  
QY 601 ATACACATCATCAATTAATTAATTAATCTGC-CGTTGGTGGCTAAGTAACTATTTGGT 659  
Db 8549 GCCATTTTCAATTTTAAAGCAATTAATCCGCAATATTTTCCGCAAAAGAAACGGTATGTT 8490





Dh 103540 CCATACCACTTTCACACGAGCTTTGCCATTCTAATAAGTAGTACGGTAC 103481  
Qy 1020 TTTTGGCATTTGACCTGTGATGGCGATGATGGTGTGTTGTTGCTTAGATATCTC 1079  
Dh 103480 TTTTACCATTTTGAACCTGTATCCCAACATTTGGCTTTGTCGTCGGCGCA----- 103429  
Qy 1080 GTGCTTGAAGTGTATCAATGACAAATTTGAACATCGCTGATCAGAGAAATGCTTGAGCTT 1139  
Dh 103428 -----GAAATTAATTCATATCGCGATTAATCTTCCATCCGCTTTAA 103388  
Qy 1140 TGGCGGCAACAATACCTGGCGTCTTGCGTTATGCGCGGCTGATGATGATCTGATCGG 1199  
Dh 103387 GTGCGGTTTGAATTTCTGGGTTTTCACGCAAGCCCTGGCTTAATAACATATATCGC 103328  
Qy 1200 CTGAGCTTAAGAGTTGATCTGTCCAGCTGCGCAAAATGACATGAGCGCTGTGGGAGTT 1259  
Dh 103327 TTTTCAAGTAACTTCTGTATTTAACTACAGTATGAAAGGATATTTTGAAGAAATT 103288  
Qy 1260 TATCTGCCAAGTGAAGATTGGGATTAACATCGATGATGATGCTTGTGACCGCTGATTA 1319  
Dh 103267 TATCAATACCAGATGAGATTTTTCGAGTATCAATCAACGAATATTAAGCTGTGGGATA 103208  
Qy 1320 CCAAAAATTCACAGCTGACAGTCCAGAGCTGCGCAACCGACAGCGCATATTTT 1377  
Dh 103207 AAGATTAATCCACACAAAGAAAGCTGTTTTCGCAAGCCCGATGATGATATATTTT 103150

RESULT 12  
BD426631.12/c  
WPCOMMENT

Sequence split into 19 fragments LOCUS BD426631 Accession BD426631

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| BD426631_00   | 1       | 110000  |
| BD426631_01   | 100001  | 210000  |
| BD426631_02   | 200001  | 310000  |
| BD426631_03   | 300001  | 410000  |
| BD426631_04   | 400001  | 510000  |
| BD426631_05   | 500001  | 610000  |
| BD426631_06   | 600001  | 710000  |
| BD426631_07   | 700001  | 810000  |
| BD426631_08   | 800001  | 910000  |
| BD426631_09   | 900001  | 1010000 |
| BD426631_10   | 1000001 | 1110000 |
| BD426631_11   | 1100001 | 1210000 |
| BD426631_12   | 1200001 | 1310000 |
| BD426631_13   | 1300001 | 1410000 |
| BD426631_14   | 1400001 | 1510000 |
| BD426631_15   | 1500001 | 1610000 |
| BD426631_16   | 1600001 | 1710000 |
| BD426631_17   | 1700001 | 1810000 |
| BD426631_18   | 1800001 | 1830121 |

Continuation (13 of 19) of BD426631 from base 1200001 (BD426631 The Nucleotide Sequence)

Query Match 11.6%; Score 160.4; DB 6; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 4.2e-28;  
Matches 663; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy 1 TAAAGGTGAACCATGTATAACAATTATACCTGCTCATATTAATTTTAAATGATC 60  
Dh 4439 TAAACCTTAGCTAATATGCTTAATTTCTGCGCGCTTTTCAAAAGAAAGCAACTGATC 4380  
Qy 61 CATGCTAGACAGACAGGTGACAAAGACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120  
Dh 4379 GAGACTTGACAGAGCGGACAAATATATCA-----TATCTCCGCTTTGCAA 4332  
Qy 121 TACTGCGCATCAAGGTACCAACTGATATATTTTGAATGCCAATGTGCATACCTT 180  
Dh 4331 TGTGGGCTGAATAAATTTATATGCTTGTTCATTTGATGAAACA-----ATACTTGG 4278  
Qy 181 GAGTAAATCTTGCTCAATTAATCGGGGATCTTGAACGATTAATAACACTATCCACCA 240  
Dh 4277 CGATGAATTTTTCGACAGCGCCACCACTCTGACCAAAACATTAACAAATATGTGTGG 4218

Qy 241 CTGCAATACATATGTCTGACAGCTCAATAATCTTGACCTTTACAGAGCCACCAAAAT 300  
Dh 4217 TGTATTAATTAATTCAGCTAATTTTCAAAAATACGCCCTTTTCCGTCCGGCTAGCA 4158  
Qy 301 CAATGCAAGTATGATATATGCAATTAACGTGCCAAGCCCAATCAATTTGCTGCAATGGTTGA 360  
Dh 4157 CAATGCAATTTAACCTCAATATA-----AAGCCAGCCAAATGCGCAACGTCTACT 4107  
Qy 361 GCGAGTTTGTATACCTTTTGAATCAATTAATATAGCTCTGATATCAATATCAATATA 420  
Dh 4106 CCCCAATTTTGTGCTTTAGATCATTAATCCAGAAATGCCATTTAGCTTATGCACTAA 4047  
Qy 421 TTCACAGCATGCGGTAGCCCTTTGAAGCTTAAGGGATCAAGCATGCGATCCATGGG 480  
Dh 4046 TTGAATAACATGATCTTAACCTTTTGAATGACGAAGTGGGTATGCAATTTGAATCTAAAT 3987  
Qy 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTCAAGATTAAGCATTTGATGATATGCTTGGC 540  
Dh 3986 AATACCTATAGCTGTGCGCAATGCTGTGCTGCCAAATGTTCATATTAATTTATGCGCAC 3927  
Qy 541 TTTGATTTTAAATGATGATTTGTAAGTTGTTTTTTGTGATATATATGCAATACC 600  
Dh 3926 AACCAATGATGCTTTCTTCAAGAGTAAATACCTTCATCTTTTACATTAATATGCTT 3867  
Qy 601 ATCACCATCATCAATAAATAAATAATCTGC-CGTTTGGTGGCTAAGTATGATTTGTTG 659  
Dh 3866 GCCATTTTCAGTTTATAGCCAAATATCCGCACTATTTTCCGCAAAAGAAAGGTATGTT 3807  
Qy 660 TGTATATGCTGATTTTGGGGCAAGCTTGTCAATGTGTCAAGCATTTGCTGCTAAT 719  
Dh 3806 CGCTGATTTTGTGTTTCCCAAAAGTCAGCC----- 3775  
Qy 720 TGGCATTCGTGTGATGATGATGGCGTTTGGCAATTTTCAAAAATAGCAATTTTGTG 779  
Dh 3774 ---TATCTTCAATGTTTCAACACACTACTTGAATATGATTAATGCTATTTTGTCT 3718  
Qy 780 CCAAAATATCAACATGCCATCGTGGCGATCAAGATGATCAGCGAGATTTAAATTTG 839  
Dh 3717 GCGATTAATCTTAAATCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 3658  
Qy 840 TTGCCCCGTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAACTTGAACAGCTCAAGCA 899  
Dh 3657 TCGCTGCGGACCTTTTAAGCTATATATGATGATGATGATGATGATGATGATGATGAT 3598  
Qy 900 CCACCAATCATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959  
Dh 3597 CATTAAGTTCAAACTTATCTATTC---AACATGACAAAGCGGAAATCCCAATATTTCCGC 3541  
Qy 960 CGACACCAACATATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
Dh 3540 CCATACCAACTTTCACACGACGAGCTTTCGCAATTTCAATTAAGTATGATGATGATGAT 3481  
Qy 1020 TTTTGGCATTTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
Dh 3480 TTTTACCAATTTTAACTGTAAATCCCAAAATTTGCTGTGCTGCTGCGCGCA----- 3429  
Qy 1080 GTGCTTGAAGTATCAATGAACATTTGAACATGCTGATCAAGAAATGCTTGAGCTT 1139  
Dh 3428 -----GAAATTAATTCATATCGCGATTAATCTTCCATCCCGCTTTAA 3388  
Qy 1140 TGGCGGCAACAATACCTGGCTTCTTGAGTTAATCGGGGCTGATGATGATCTGATCGG 1199  
Dh 3387 GTGCGGTTTGAATTTCTGGGTTTTCACGCAAGCCCTGGGCTTAATAACATCATATCGC 3328  
Qy 1200 CTGAGCTTAAGATTCACTGTCCAAAGCTGCCAAATGACATGAACGCTGTGGGCAATT 1259  
Dh 3327 TTTCAAGTAACCAATTCCTGATTTAACTACAGTATGAAGAGGATATTTTGAAGAAATT 3268  
Qy 1260 TATCTGCCAAGTGAAGATTGGGATTAACATCGATGATGATGATGATGATGATGATGAT 1319  
Dh 3267 TATCAATACAGTATGATTTTTCGAGTATCAATCAACGAATATTAATGCTGTGGGATA 3208  
Qy 1320 CCAAAAATTCACAGCTGACAGTCCAGAGCTGCCAAACGACGAGCGCATATTTT 1377

Db 3207 AGAGATATATCCACAGAAAGACCTGTTTGGCAAGCCCGATGATATATTTT 3150

RESULT 13  
AR274513\_11/c  
WPCOMMENT

Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| AR274513_00   | 1       | 110000  |
| AR274513_01   | 100001  | 210000  |
| AR274513_02   | 200001  | 310000  |
| AR274513_03   | 300001  | 410000  |
| AR274513_04   | 400001  | 510000  |
| AR274513_05   | 500001  | 610000  |
| AR274513_06   | 600001  | 710000  |
| AR274513_07   | 700001  | 810000  |
| AR274513_08   | 800001  | 910000  |
| AR274513_09   | 900001  | 1010000 |
| AR274513_10   | 1000001 | 1110000 |
| AR274513_11   | 1100001 | 1210000 |
| AR274513_12   | 1200001 | 1310000 |
| AR274513_13   | 1300001 | 1410000 |
| AR274513_14   | 1400001 | 1510000 |
| AR274513_15   | 1500001 | 1610000 |
| AR274513_16   | 1600001 | 1710000 |
| AR274513_17   | 1700001 | 1810000 |
| AR274513_18   | 1800001 | 1830121 |

Continuation (12 of 19) of AR274513 from base 1100001 (AR274513 Sequence 1 from patent U

Query Match 11.6%; Score 160.4; DB 6; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 4.2e-28;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGCGTGAACCACTGTAACAACCTTATGACCTGCTCATTAATTTTAAATGATC 60  
DB 104439 TAAAGCTGAGCTAAATGCGCTAAATCTTGCGCGCTTTTCAAAAGAACCAATGATC 104380  
QY 61 CAGCTAGACAGACAGGAGTGAACAAAGACAGCCCTAGCGCTTGAAGCTGCGCA 120  
DB 104379 GAACTTGAACAGAGGAGGACCAATATATCA-----TATCTCCGCTTGGCA 104332  
QY 121 TACTGCCGATCAAGTGTACCACTGATATATTTAGATGACCAATGTGACCTT 180  
DB 104331 TGTGGGCGTAAATATCTATCGCTTGTTCATTTGATGCAACA-----ATTAATCTTG 104278  
QY 181 GAGTAATCTTGTCTCAATTAATGAGGAGCTTGAACGATTAATACACTATCCACA 240  
DB 104277 CGATGAAATTTTGAAGAGGAGCCCATCTCGACCAAAACAATTAATATGTGG 104218  
QY 241 CTGATTAACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT 300  
DB 104217 TTGATTAATTAATTCAGCTAATCTGAAATATGAGCCCTTTCCTCCGCTGACCA 104158  
QY 301 CAATGCCAGTATGATATGCAATTAACCGTGCAGAGCCCATCAATGTGCTCAATGTTGA 360  
DB 104157 CAATGCAATTAACCTCATATTA-----AAGCCGCAATGACAGCACTGATC 104107  
QY 361 GCCGATGTTTGAACCTTTGAATCAATTAATATGCTTGTATCAATATCATATATA 420  
DB 104106 CCCCAATTTTGTCTTGAAGTCAATTAACCAAGATGCAATTAATGATGCACTAA 104047  
QY 421 TTCACAGCAGATGGGATAGCCCTTTGAAGTCTTAAGGATTCAGAGATGCAATGCGG 480  
DB 104046 TTGAAACAGATGATCTTAACCTTTGAAGATGCAAGATGCGGTGAATTAATCTAAAT 103987  
QY 481 AAGATTGCGAGCTGTGCTTAAGGCAAGTGAAGTGAAGGCAATGATGATGATGCTTGGC 540  
DB 103986 AATACCTAATAGCTGTGCGCAATGCTGTGCTGCCAAATGTCATTAATTAATGCGGAC 103927  
QY 541 TTTGATTTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 103926 AACCAATGATGCTTCTTCAACAGGTAATATCACTCATCTTTTACATTAATATTTGCTT 103867

QY 601 ATGACCATCATCAATTAATTAATTAATCTGC-CGTTTGTGCTAATGTAATTTGTTG 659

DB 103866 GCCATTTTCAGTTTATGACCAATATCCGCAATATTTTCCGCAAAAGAAACGGATGTTT 103807

QY 660 TGATTAATGATGATTTGGGGCAACGCTTGTCAAGTGTGCAAGCATTTGCTGGCTAAAT 719

DB 103806 CGCTTGATTTTGTCTTTTCCCAAAAGTCAGC----- 103775

QY 720 TGGCATGCTGTTGTGATGATGATGCGGTTTGGCAATTTTCAAAATACGAATTTTGTG 779

DB 103774 ---TATCTCATTTGTTCAACACACTTACCATTAATATTAATGCGTAATTTGCTT 103718

QY 780 CCAATTAATCCACAGTCCATCGTGCAGTCAAGATGATGAGGAGAGATTTAAATTTG 839

DB 103717 GCGATTAATCTTTAAATCAATTAATGATGATGATGATGATGATGATGATGATG 103658

QY 840 TTGCCCCCTGAGCTGCAATTTGAGATATGCTCAAGCTGAAACTTGAACAGCTCAAGCA 899

DB 103657 TCGCTGCGCAGCTTTTAACTATGATGATGATGATGATGATGATGATGATGATGATG 103598

QY 900 CCAACCAATTCATATTTTGGATGCTTAATTAATTAATTAATTAATTAATTAATTAAT 959

DB 103597 CATTAAGTTCAACATCTTCAATC---AACATGACAAAGGCGGATCCCAATATTTCCG 103541

QY 960 CGACACCAACATCATGCTGCAATGTTTGGCAATCTGCTACTAATGCTGAACAGTGC 1019

DB 103540 CCAATCAACTTTTCAACAGCAGCTTTCGCAATTTCAATTAATTAATTAATTAATTAAT 103481

QY 1020 TTTTGGCATTTGAGCTGATGATGCGATATGCTGTTGTTGTTGCTTGAATAGATCTC 1079

DB 103480 TTTTACATTTTGAACCTGTAATCCCAACATTTGCTTGTGCTGCGCGCA----- 103429

QY 1080 GTGCTTGAAGTATCAATGAACATGTAACATGCTGATCAAGAAATGCTTGAAGCTT 1139

DB 103428 -----GAATTAATTAATATCGCGATATCTTCCATCCCGCTTAA 103388

QY 1140 TGGCGGCAACATCTGCGCTTCTTGGGTTAATGCGGGGCTGATGATGATGATGATGAT 1199

DB 103387 GTGCGGTTGAATTTTGTGTTTAAACGCAAGCCCTGGGCTTAATTAATTAATTAAT 103328

QY 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259

DB 103327 TTTTCAATTAACATTTCTGATTTAACTACAGTATGAAGAGGATATTTTGAAGAT 103268

QY 1260 TATCTGCCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319

DB 103267 TATCATTAACAGTATGATTTTTCAGATATCAATCAACGAATATTAATGCTGTTG 103208

QY 1320 CCAAAATTTCAAGCTGACAGTCCAGAGCTGCCAAACCGACGCAATATTTT 1377

DB 103207 AGAGATTAATCCACAGAAAGACCTGTTTGCAGAGCCGATGATGATTAATTTT 103150

RESULT 14  
AR274513\_12/c  
WPCOMMENT

Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| AR274513_00   | 1       | 110000  |
| AR274513_01   | 100001  | 210000  |
| AR274513_02   | 200001  | 310000  |
| AR274513_03   | 300001  | 410000  |
| AR274513_04   | 400001  | 510000  |
| AR274513_05   | 500001  | 610000  |
| AR274513_06   | 600001  | 710000  |
| AR274513_07   | 700001  | 810000  |
| AR274513_08   | 800001  | 910000  |
| AR274513_09   | 900001  | 1010000 |
| AR274513_10   | 1000001 | 1110000 |
| AR274513_11   | 1100001 | 1210000 |
| AR274513_12   | 1200001 | 1310000 |
| AR274513_13   | 1300001 | 1410000 |

AR274513\_14 1400001 1510000  
AR274513\_15 1500001 1610000  
AR274513\_16 1600001 1710000  
AR274513\_17 1700001 1810000  
AR274513\_18 1800001 1830121  
Continuation (13 of 19) of AR274513 from base 1200001 (AR274513 Sequence 1 from patent U

Query Match 11.6%; Score 160.4; DB 6; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 4.2e-28;  
Matches 663; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGCGTGAACCATGTAAACAACCTTATCACTCGCTCAATTAATTTTAATTGATC 60  
DB 4439 TAAACCTTAGCTTAATGCGTAATCTTCGCGCGCTTTCAAAAGAACCAACGATC 4380  
QY 61 CATGTAGACCAAGAGGTGACAAAGACAGCCCTTAGCGGCTTTGACTGCGCA 120  
DB 4379 GAGACTTGACACAGCGGACCAATATACCA-----TATCTCGCTTGCA 4332  
QY 121 TACTGCCGATCAAGTGAACAACCTGATATATTTAGATGCCAAATGTGCATCACCTT 180  
DB 4331 TGTGGGCGTAAATAATTTCTATGCTTGTTCATTTGATGAAACA-----ATACTTGG 4278  
QY 181 GAGTAATCTGTCTCAATTAATGGGGCATCTTGAACGATTAATACACTATCCACA 240  
DB 4277 CGATGAAATTTTGCAAGCAGCGCACCATCTCGACAAACAAATTAATATGTGTGG 4218  
QY 241 CTGCATACATATGTGACAGCTCAGTAAATCTTGACCTTTACAAAGCCCAACCAAT 300  
DB 4217 TTGATTAATTAATTCAGCTAATTTCTGAAATAATGACCCCTTTCGCTCGCTAGCA 4158  
QY 301 CAATGCCAGTATGATATGCAATTAACCGTGCACAGCCCATCAATTTGCTGCATGTTGA 360  
DB 4157 CAATGCAATTAACCTCAATATA-----AAAGCCAGCAATGACGAACTGACT 4107  
QY 361 GCGGATGTTTGAACCTTTTGAATCAATTAATATGCTTGTCTATCAATATCAATATA 420  
DB 4106 CCCCAATTTGCTTTAGAGTCATTAATCCAAAGCAAGCCATTTGCTGATGCACTAA 4047  
QY 421 TTCACAGCAGCGGTAGCCCTTTGAAAGCTTAAGGATTAAGAGATGAGCATCCATGG 480  
DB 4046 TTGAAACCATGATCTTAACCTTTGAAAGAGCAAGGTGGGATGCAATTTGAATTAAT 3987  
QY 481 AAGATTGGAGCTGTGCTTAAGGCAAGTGAAGTAAGGATGAGTATGCTTGGC 540  
DB 3986 AATACCTATAGCTGTGCAAGCTGTGTGCAAAATGTTCAATTAATTAATGCGGAC 3927  
QY 541 TTTGATTTTAATGATGATGATGTAAGTTGTTTGTGATGATTAATGCAATACC 600  
DB 3926 AACCAATGATGCTTCTTCAAGAGTAATCACTTCATCTTTTACCAATTAATATGCTT 3867  
QY 601 ATCACCATCATCAATTAATTAATAATCTGC-CGTTTGGGTGCTAAGTATTTGTTGG 659  
DB 3866 GCCATTTTCAGTTTTHAGCAATTAATCCGACATATTTCCGCAAAAGAAACGATATGTT 3807  
QY 660 TGAATATGATGATTTGGGGCAACGCTTGTCAAGTGTGTCAAGCATGCTTGGCTAAT 719  
DB 3806 CCTTGATTTTGGTTTCCCAAAAGTCAAGC----- 3775  
QY 720 TGGCATCTGCTTGTGATGATGATGAGCGTTTGGCAATTTTCAAAATATCCGATTTTGTG 779  
DB 3774 ---TATCTTCATTTGTTCAACACCTTACCTTTAGCATTAATGAAGCGTAATTTTGGCTT 3718  
QY 780 CCAATTAATTCACATGCGCATGTCGCAATCAAGATGATCAAGGAGAGATTAAATTTG 839  
DB 3717 GCGGATTAATCTTAAATCAATTAAGCATTCATGATGATCTTCAAGCAATTCAGACAG 3658  
QY 840 TTGCGCTTGAAGTGCCTTGAATTTGATATGCTCAAGTGAATTTGAACGCTCAAGCA 899  
DB 3657 TCGCTGCGCAGCTTTTAAGCTTAATGCTTGTGTCAAGCTGAAGAACTGAAGTCTTAAGTA 3598  
QY 900 CACCAATTCATATTTTGGATGTTAATTAATTCAAAGTCAAGCGCTGCAATTAATACCG 959

DB 3597 CATTAAGTTCACAAATCTTCAATC---AACATGACAAAGCGGAAATCCCAATATTTCCCG 3541  
QY 960 CGACACCAACATTCATGCTGATGTTTGGCATCTCGCTCACTAATGTGTAACGTGC 1019  
DB 3540 CCAATCAACCTTTCACACAGCAGCTTCCGATTCATTAACATTAAGTAGTACGATAC 3481  
QY 1020 TTTTGGCATTTTGAAGCTGTGATGCGCATGATTTGATGTTTGTGCTTGAATGATCTC 1079  
DB 3480 TTTTACATTTTGAACCTGTAAATCCCAATTTGGCTTTTCGCTGCGGCA----- 3429  
QY 1080 GTGCTTGAAGTATCAATGAACCAATTTGAACCTGTATCAACAGAAATGCTTGAAGCTT 1139  
DB 3428 -----GAATTAATTCATATGCGGATTAATCTTCCATCCCGCTTTAA 3388  
QY 1140 TGGCGGCAACATTAATCTGCGCTTCTTGGGTATGCGGAGGCTGATGATCTGATCGG 1199  
DB 3387 GTGCGGTTGAATTTCTGTGTTTTCACGCAAGCCCTTGGCTTAATACATTAATTCGC 3328  
QY 1200 CTGAGCTTAAGATTCATCTGTCAGAGCTGCCAAATGACAAATGAACGCTGTGGCAGTT 1259  
DB 3327 TTTCAAGTAACCATCTCTGATTTAACTACAGTATGAAGAGGATATTTGAGGAAGTT 3268  
QY 1260 TATCTGCCAAGTGAATTTGAGATTAATCATGATGATGCTAATTTGTAACCGTATTA 1319  
DB 3267 TATCAATACCAATGAGATTTTTCGAGTATCAATCAAGAAATATTAAGCTGTGGGATA 3208  
QY 1320 CCAAAATAATTCACAGCTGACAGTCCAGAGCTGCCCAACGAGCGGATATTTT 1377  
DB 3207 AAGATTAATTCACACAAAGAAAGCTTTTTCAGAGCCGATGATTAATTAATTTT 3150

# RESULT 15

AR632719\_11/c

WPCOMMENT

Sequence split into 19 fragments LOCUS AR632719 Accession AR632719

| Fragment Name | Begin   | End     |
|---------------|---------|---------|
| AR632719_00   | 1       | 110000  |
| AR632719_01   | 100001  | 210000  |
| AR632719_02   | 200001  | 310000  |
| AR632719_03   | 300001  | 410000  |
| AR632719_04   | 400001  | 510000  |
| AR632719_05   | 500001  | 610000  |
| AR632719_06   | 600001  | 710000  |
| AR632719_07   | 700001  | 810000  |
| AR632719_08   | 800001  | 910000  |
| AR632719_09   | 900001  | 1010000 |
| AR632719_10   | 1000001 | 1110000 |
| AR632719_11   | 1100001 | 1210000 |
| AR632719_12   | 1200001 | 1310000 |
| AR632719_13   | 1300001 | 1410000 |
| AR632719_14   | 1400001 | 1510000 |
| AR632719_15   | 1500001 | 1610000 |
| AR632719_16   | 1600001 | 1710000 |
| AR632719_17   | 1700001 | 1810000 |
| AR632719_18   | 1800001 | 1830121 |

Continuation (12 of 19) of AR632719 from base 1100001 (AR632719 Sequence 1 from patent U

Query Match 11.6%; Score 160.4; DB 6; Length 110000;  
Best Local Similarity 49.6%; Pred. No. 4.2e-28;  
Matches 663; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGCGTGAACCATGTAAACAACCTTATCACTCGCTCAATTAATTTTAATTGATC 60  
DB 104439 TAAACCTTAGCTTAATGCGTAATCTTCGCGCGCTTTCAAAAGAACCAACGATC 104380  
QY 61 CATGTAGACCAAGAGGTGACAAAGACAGCCCTTAGCGGCTTTGAATGCGTGC 120  
DB 104379 GAGACTTGACACAGCGGACCAATATACCA-----TATCTCGCTTGCA 104332  
QY 121 TACTGCCGATCAAGTGAACCAACCTGATATATTTTAATGCAATGATGATCACCTT 180  
DB 104331 TGTGGGCGTAAATAATTTCTATGCTTGTTCATTTGCAATGCAACA-----ATACTTGG 104278

QY 181 GAGTAATCTTGTCTCAATTATCGGGGCACTTGAACGATTAATAACACACTATCCACCA 240  
Db 104277 CGATGAAATTTTTCAGACAGCGGCACTCGACCAAAACATTAACAAATATATGTGG 104218  
QY 241 CTGATTAACATATCTGTACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT 300  
Db 104217 TTGTTAATTAATTAATCTGCTAATTTTGAAAAATCAGCCCTTTTCCTCTCGCTACCA 104158  
QY 301 CAATGCCAGTATGATATGCAATTAACCGTSCCAAGCCCATTAATGCTGCAATGTGTA 360  
Db 104157 CAATGCAATTTTACCCCAATTA-----AAGCCAGCCCAATGACGAACTGTACT 104107  
QY 361 GCCGATCTTGTATCTTTTGAATCAATAAATATGCTTGTATCAATATCATATATA 420  
Db 104106 CCCACATTTTGTCTTTAGTGTATTAATCAACGAATGCCATTAGCTGATGACCTAA 104047  
QY 421 TTGACAGGATGCGGTAGCCCTTTGAAAGCTTAAGGATCAAGCATGGCATCCATGG 480  
Db 104046 TTGAAAACGATGATCTAACCTTTGAAATGAAGAGTCCGCTACGAATGTAAAT 103987  
QY 481 AAGATTGCGAGCTGTGCTTGAAGCAAGTGCAGATAGGCATTTGATGATGCTTGGC 540  
Db 103986 AATACCTATAGCTTGTGCAATGTCTGTGCTGCCAAATGTTCAATATATATGCGACC 103927  
QY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 103926 ACCAATGATGCTCTTCAAGATTAATACATCTTATCATTAATATATGCTT 103867  
QY 601 ATCAACATCATCAAAATTAATAAATCTGC-CGTTGCTGCTAAGTAACTTTGTTG 659  
Db 103866 GCCATTTTCAGTTTTCAGCAATATCCGCACTATTTTCGCAAAAGAAACGATATGTT 103807  
QY 660 TGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719  
Db 103806 CGCTTGAATTTTCGTTTCCCAAAAGTACG----- 103775  
QY 720 TGGCATGCTGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 779  
Db 103774 ---TATCTTCAATGTTCAACACACTTACTTTAGCATTAATGATATGATATTTTGTCT 103718  
QY 780 CCAATATATCACCAATGCCATCGTCCGATCAAGATGATCAAGAGATTTAAATG 839  
Db 103717 GGGATATATCTTCAATCAATATGATGATGATGATGATGATGATGATGATGAT 103658  
QY 840 TTGCCCCGTGAGCGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 899  
Db 103657 TCGCTGCCGAGCTTTTAACTATATGATGATGATGATGATGATGATGATGAT 103598  
QY 900 CCAACCAATCATATTTTGGATGATGATGATGATGATGATGATGATGATGATGAT 959  
Db 103597 CATTAAGTTCAATCTTCAATC---AACATGACAAAGCGGAAATCCCAATATTTCCG 103541  
QY 960 CCAACCAACATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
Db 103540 CCATACCACTTTCAACAGAGCTTTGCAATTTCAATACTAAAGTAAAGGTAC 103481  
QY 1020 TTTTGGATTTGAGCGCTGTATGAGCGATGATGATGATGATGATGATGATGATGAT 1079  
Db 103480 TTTTACCATTTGAACCTGTAATCCCAATTTGCTTTGCTGCGCGGCA----- 103429  
QY 1080 GTGCTTGAAGTATCAATGAACATGAATGATGATGATGATGATGATGATGATGAT 1139  
Db 103428 -----GATTAATTAATTAATGCGGATTAATCTTCCATCCGCTTTAA 103388  
QY 1140 TGGCGGCGCAATACTGCGGCTTGTGATTAATGCGGCGCTGATGATGATGATGATGAT 1199  
Db 103387 GTGCGGTTGATTTTGTGTTTTCGCAAGCCCTGGGCTAATTAACATCATATCGC 103328  
QY 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
Db 103327 TTTCAAGTAACTATTTCTGATTTAACTACAGATGATGATGATGATGATGATGAT 103268  
QY 1260 TATCTGCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319

Db 103267 TATCAATACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 103208  
QY 1320 CCAAAATTTCAAGCTGACAGCTGACAGCTGCGCCAAACGAGAGCGCATATTTT 1377  
Db 103207 AGAGATATCCACACAAAGAAAGACTGTTTTCGCAAGCCGATGATGATATATTT 103150

Search completed: May 12, 2006, 09:55:03  
Job time : 7232 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 07:54:39 ; Search time 6041 Seconds  
(without alignments)  
10687.994 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380  
Sequence: 1 taagacgctgaccatgttaa.....acgacggcatatttttga 1380

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_g861:\*  
10: gb\_g882:\*  
11: gb\_g883:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| C 1        | 98.2  | 7.1         | 817    | 10    | CZ543630 SRAA-aad5  |
| C 2        | 92.8  | 6.7         | 709    | 6     | CF306940 HD11-05-   |
| C 3        | 92    | 6.7         | 892    | 9     | B2579116 msh2.6205  |
| C 4        | 77.2  | 5.6         | 555    | 10    | AG943479 Drosophila |
| C 5        | 69.8  | 5.1         | 1445   | 9     | B2568959 pasc2-164  |
| C 6        | 62.2  | 4.5         | 623    | 10    | AG952088 Drosophila |
| C 7        | 56.8  | 4.1         | 628    | 10    | CL681264 PRI0156a   |
| C 8        | 55    | 4.0         | 474    | 8     | DR032465 B2559806   |
| C 9        | 52.6  | 3.8         | 1171   | 9     | B2559806 pasc2-164  |
| C 10       | 50.2  | 3.6         | 838    | 10    | CZ545153 SRAA-aad6  |
| C 11       | 49.8  | 3.6         | 572    | 9     | BH795426 BHM03766   |
| C 12       | 49.4  | 3.6         | 452    | 2     | BE184674 PM4-HT068  |
| C 13       | 47.6  | 3.4         | 1101   | 10    | CNS00000            |
| C 14       | 47.2  | 3.4         | 1101   | 10    | CNS00000            |
| C 15       | 46    | 3.3         | 823    | 8     | DR505933 WS02715.B  |
| C 16       | 45    | 3.3         | 590    | 1     | AU287701 AU287701   |
| C 17       | 45    | 3.3         | 757    | 8     | DR473894 WS00960.B  |
| C 18       | 45    | 3.3         | 907    | 8     | DN922769 43863.2.C  |
| C 19       | 44.4  | 3.2         | 442    | 8     | DR565020 WS02632.C  |
| C 20       | 44.4  | 3.2         | 490    | 7     | CO253110 WS00818.B  |
| C 21       | 44.4  | 3.2         | 794    | 8     | DR473810 WS00960.B  |
| C 22       | 43.6  | 3.2         | 692    | 9     | BH983612 odes20g11. |

|      |      |     |      |    |           |                     |
|------|------|-----|------|----|-----------|---------------------|
| 23   | 43   | 3.1 | 1035 | 10 | CNS00025  | AL097523 Drosophila |
| 24   | 42.8 | 3.1 | 1101 | 10 | CNS0100X  | AL098379 Drosophila |
| C 25 | 42.4 | 3.1 | 1101 | 10 | CNS0039G  | AL063921 Drosophila |
| C 26 | 42.2 | 3.1 | 747  | 10 | C2688937  | C2688937 OM_Ba022   |
| C 27 | 42.2 | 3.1 | 915  | 9  | AZ543308  | AZ543308 ENTGK93TR  |
| C 28 | 42   | 3.0 | 941  | 11 | CNS05154  | AL342317 Tetradon   |
| C 29 | 42   | 3.0 | 1022 | 11 | CNS070NB  | AL436797 T3 end of  |
| C 30 | 41.8 | 3.0 | 517  | 1  | AU287702  | AU287702 AU287702   |
| C 31 | 41.8 | 3.0 | 1101 | 10 | CNS0000D1 | AL065514 Drosophila |
| C 32 | 41.6 | 3.0 | 616  | 9  | BH383040  | BH383040 AG-ND-127  |
| C 33 | 41.6 | 3.0 | 714  | 9  | BH374901  | BH374901 AG-ND-127  |
| C 34 | 41.6 | 3.0 | 747  | 6  | CA920035  | CA920035 EST637753  |
| C 35 | 41.4 | 3.0 | 639  | 8  | CK639436  | CK639436 UCRPT02.4  |
| C 36 | 41.4 | 3.0 | 760  | 9  | BH367641  | BH367641 AG-ND-127  |
| C 37 | 41.2 | 3.0 | 378  | 9  | BH382726  | BH382726 AG-ND-127  |
| C 38 | 41.2 | 3.0 | 434  | 8  | T92809    | T92809 Ye23d06.r1   |
| C 39 | 41   | 3.0 | 563  | 1  | AM020818  | AM020818 AM020818   |
| C 40 | 41   | 3.0 | 726  | 9  | BH047969  | BH047969 RPT1-24-3  |
| C 41 | 41   | 3.0 | 902  | 9  | AZ685437  | AZ685437 ENTLE37TF  |
| C 42 | 40.8 | 3.0 | 693  | 5  | BQ446218  | BQ446218 UR-H-EU1-  |
| C 43 | 40.8 | 3.0 | 939  | 10 | CNS017ZY  | AL108712 Drosophila |
| C 44 | 40.6 | 2.9 | 635  | 8  | DN157692  | DN157692 GCN001110  |
| C 45 | 40.6 | 2.9 | 779  | 10 | CW381115  | CW381115 Feb000110  |

#### ALIGNMENTS

RESULT 1  
LOCUS CZ543630/c  
DEFINITION SRAA-aad51b03 g1 Strongyloides ratii whole genome shotgun library  
(SRAAGS 004) Strongyloides ratii genomic, genomic survey sequence.  
ACCESSION CZ543630  
VERSION CZ543630.1 GI:64673383  
KEYWORDS GSS.  
SOURCE Strongyloides ratii  
ORGANISM Strongyloides ratii  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Panagrolaimidae; Strongyloidea; Strongyloidea.  
REFERENCE 1 (bases 1 to 817)  
Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Pape,D.,  
Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,  
Clifton,S.W. and Wilson,R.  
Genome Survey sequences from the rat parasitic nematode  
Strongyloides ratii  
Unpublished (2005)  
JOURNAL Contact: Mitreva M  
COMMENT Washington University in St. Louis  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: nematode@wustl.edu  
Genomic DNA was provided by Fiona Thompson  
(F.Thompson@bristol.ac.uk) and Mark Viney  
(Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,  
UK.  
Class: shotgun.  
Location/Qualifiers  
1..817  
/organism="Strongyloides ratii"  
/mol\_type="genomic DNA"  
/strain="Isotemale line ED321 heterogonic"  
/db\_xref="taxon:34506"  
/dev\_stage="infective larval stage (L3)"  
/lab\_host="GSI0"  
/clone\_lib="Strongyloides ratii whole genome shotgun  
library (SRAAGS 004)"  
/note="Vector: pOTW13; Site\_1: BstXI; Site\_2: BstXI;  
Strongyloides ratii genomic DNA was randomly sheared,  
end-repaired and size fractionated to enrich for 2-4 kb  
fragments. Genomic DNA was provided by Fiona Thompson





ORIGIN /note="clinical isolate 2-164 whole genomic shotgun library."

Query Match 5.1%; Score 69.8; DB 9; Length 1445;

Best Local Similarity 55.0%; Pred. No. 2.5e-08; Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 327 CCGTGCACAGCCCATCAATGCTGCAATGCTGAGCCGATGTTGATCTTTGATCAT 386  
| | | | |  
DB 87 CGGACACCCAGCCCTCGATCGCCGCCAGGCGGCGCGGACGTTGCTGGAATCT 146  
| | | | |  
QY 387 TAAATATGCTCTGATCATATATCATATATATTCACAGCATGCGGTAGCCCTTTGA 446  
| | | | |  
DB 147 CGTAGAGCTTCAGCCCTCGCTCGCTGCCATCCCATGCGCATGACGACGCGGAAA 206  
| | | | |  
QY 447 AAGCTTAAGGGATCAAGCATGGCATCATATGGAATGGACATGCTGCTTAAGGCA 506  
| | | | |  
DB 207 ACACCTTCAGCGCGCCAGCATGCGCTGAAACGACAGCGCCGACATGGCCACGCCCA 266  
| | | | |  
QY 507 GTGCAGATMAAGCATGAGTAGTATGCTTGTCTTGTATTTTAAATGATGATGGTA 566  
| | | | |  
DB 267 GCGCGGAGACGGGCTGGAATAGTTGTGGGCGCACGATCTTCAGTTGCCCAACGCCCA 326  
| | | | |  
QY 567 AAGTTTGT 575  
| | | | |  
DB 327 GCAGCTTGT 335  
| | | | |

RESULT 6 AG952088 623 bp DNA linear GSS 01-FEB-2005

LOCUS Drosophila sechellia DNA, clone: DSE1-004A05.F.fa, genomic survey

DEFINITION sequence.

ACCESSION AG952088

VERSION AG952088.1 GI:58473786

KEYWORDS GSS.

SOURCE Drosophila sechellia

ORGANISM Drosophila sechellia

REFERENCE Hattori, M., Toyoda, A., Murakami, K., Kuroki, Y., Fujiyama, A.,

AUTHORS Toshiro, T.K. and Sakaki, Y.

TITLE BAC end sequences of library DNB1

JOURNAL Unpublished

2 (bases 1 to 623)

Hattori, M.

Direct Submission

Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the BAC library DSE1

For BAC library availability, please contact Masa-Toshi Yamamoto

(yamamoto@kit.jp).

Submitted (30-11-2004) by Masahira Hattori,

RIKEN, Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,

Fax:81-45-503-9170)

This work was done in collaboration with Yamamoto, M-T. Drosophila

Genetic Resource Center

Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan

Tel: 81-75-873-2660 FAX: 81-75-861-0881

PRIMERS

Sequencing : F

LIBRARY

Vector : PKS150

R Site 1 : SacI

R Site 2 : SacI.

FEATURES Location/Qualifiers

source

1..623

/organism="Drosophila sechellia"

/mol\_type="genomic DNA"

/db\_xref="taxon:7238"

/clone="DSE1-004A05.F.fa"

/clone\_1ib="DSE1 Drosophila BAC library"

ORIGIN

Query Match 4.5%; Score 62.2; DB 10; Length 623;

Best Local Similarity 51.0%; Pred. No. 3.1e-06; Matches 176; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

QY 703 GCATTCCTGGCTAAATGAGCATGCTGCTGATAGATGAGCGGTGGCAATTTTCA 762  
| | | | |  
DB 285 GATATATCAGCAGATATCTCATGTGCATCTATATGAGCAACTCATACCTTTAT 344  
| | | | |  
QY 763 AATACCAATTTTGTGCAATATATCCATGCGCATGCGCATCAAGATGATCAGC 822  
| | | | |  
DB 345 CAGTTTAAATTTAGTGTATGTATGATTTCTCCATCTTCATCTATGTGATCTG 404  
| | | | |  
QY 823 GGAGATTTAAATTTGCTGCTGAGCTGCAATTTGATATGCTCAAGCTGAA 882  
| | | | |  
DB 405 TGTATATGAGCAGTGTCTGAAATGCTCCAGTTAATTTATTCATCACTCCAAATGAAA 464  
| | | | |  
QY 883 ACTTGACAGCTCAAGCAGCAGCAATTCATATTTTGGATGTTAATTAATGAGCAG 942  
| | | | |  
DB 465 AGAGAGAGTGTGATTAATTAATTTTGTGATCTC-----TTTCTAGATCCAAACAG 518  
| | | | |  
QY 943 CGTCCCAATTTTACCGCGCAGCAGCAATATCATGCTGATGTTTGCATCTCGCTAC 1002  
| | | | |  
DB 519 AACGCTTAATTTTCCACCAATAGCTACTTTTCCCTGCGAGATTTAATATGCGCTAT 578  
| | | | |  
QY 1003 TAATGCTGACAGTGTCTTGGCAATTTGAGCTGTGATGAGAT 1047  
| | | | |  
DB 579 TAGTATGATGATGTTGATTTACCATTTTCTCTGAGCCTAT 623  
| | | | |

RESULT 7

LOCUS CL691264/c

DEFINITION

CL691264

ACCESSION CL691264

VERSION CL691264.1 GI:50213172

KEYWORDS GSS.

SOURCE Pristionchus pacificus

ORGANISM Pristionchus pacificus

REFERENCE Hattori, M., Toyoda, A., Murakami, K., Kuroki, Y., Fujiyama, A.,

AUTHORS Toshiro, T.K. and Sakaki, Y.

TITLE BAC end sequences of library DNB1

JOURNAL Unpublished

2 (bases 1 to 628)

Hattori, M.

Direct Submission

Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the BAC library DSE1

For BAC library availability, please contact Masa-Toshi Yamamoto

(yamamoto@kit.jp).

Submitted (30-11-2004) by Masahira Hattori,

RIKEN, Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,

Fax:81-45-503-9170)

This work was done in collaboration with Yamamoto, M-T. Drosophila

Genetic Resource Center

Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan

Tel: 81-75-873-2660 FAX: 81-75-861-0881

PRIMERS

Sequencing : F

LIBRARY

Vector : PKS150

R Site 1 : SacI

R Site 2 : SacI.

/db xref="taxon:54126".  
/clone lib="Mixed stage foamid library of P. pacificus  
var. California"  
/note="Vector: pepifos-5 Fosmid vector"

## ORIGIN

Query Match 4.1%; Score 56.8; DB 10; Length 628;  
Best Local Similarity 48.0%; Pred. No. 0.00011;  
Matches 246; Conservative 0; Mismatches 242; Indels 24; Gaps 2;

QY 1 TAAAGCGTGAACCATGTTAACTTAATCATACCTCGCTCATTAATTTTAAATGATC 60  
DB 494 TAACCTCTTGGCCAGACGCGCAACTCATGCTCGTGTTCAGAAAGTTCTGAACCTATC 435  
QY 61 CATGCTAGCAACAGAGGTGACAAAGACAGCCCTGAGCGGTGTTGACTGAGTGCCAA 120  
DB 434 AAGGCTGGCAAGGCTGGGAGAGACAAACAT-----ATGCCCCGCGTG 390  
QY 121 TACTGCCGATCAAGTGTACCAACTGATATATTTAGATGCCAATGTGCATCACCTTT 180  
DB 389 AACACGCGAGCAGCAAGGCGCATCGCTGTTCATAGTTTCGGTTTGTCTGCACTTC 330  
QY 181 GAGTAATCTTGTCTCATTTATCGGGGCACTTTGACGATTAATACACTATTCACCCA 240  
DB 329 CGGCGCGTACGCGCCGCTGCGCGCGTCAAGCAAGAAATACAGAGTACGTTATC 270  
QY 241 CTGCAATTAACATATGCTGACAGCTCACTAAATCTTACCTTTACCAAGCCACCCAAAT 300  
DB 269 GCCATTAGGTACGCGCCAGTGGCTTAAGTCCGCGATTTACCATGCGCACCGACAA 210  
QY 301 CAATGCCAGTATGATATGCGCATTAACCGTGCACAGCCCATCAATGCTGCAATGGTTGA 360  
DB 209 CAATGAGTGTGCGCTT-----CTAGTGAAGCCATTCAGCGCCGCTTCGTAAT 159  
QY 361 GCCGATGTTGTATCTTTGATCATTAATATGCTTCTGATCAATATCATATATA 420  
DB 158 GCCGAGTGTGTGCTTTCATATCGTTAATCAACGATACCGCTGTGCTCAGACAAC 99  
QY 421 TTCACAGCATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGATCCATGG 480  
DB 98 TTCAAAGCATGCGGCGACACCATGTAATGTGTTAACGCTTTCAGGCTGTCGACGCGG 39  
QY 481 AAGATTGCGAGCTGTGCTTAAGCAAGTGCAG 512  
DB 38 TAACCTTGAGCATGCTGCGCGCGCGCG 7

RESULT 8  
DRO32465/c 474 bp mRNA linear EST 31-MAY-2005  
LOCUS bda020270D15.ab1\_061 Antrodia cinnamomea cDNA library Antrodia  
DEFINITION cinnamomea cDNA, mRNA sequence.

ACCESSION DRO32465  
VERSION DRO32465.1 GI:66834715  
KEYWORDS EST.  
SOURCE Antrodia cinnamomea  
ORGANISM Antrodia cinnamomea

REFERENCE 1 (bases 1 to 474)  
AUTHORS Tzean,S.S., Hwang,S.T., Song,H.C., Liou,R.F., Yeh,H.H. and Shiao,M.S.  
TITLE Analysis of expressed sequence tags from Antrodia cinnamomea, a polypore causing brown rot of Cinnamomum kanehira  
JOURNAL Contact: Tzean SS  
COMMENT Department of Plant Pathology and Microbiology  
National Taiwan University  
NO.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan  
Tel: 886 2 3366 4595  
Fax: 886 2 2362 0639  
Email: sctc@ntu.edu.tw

PCR Primers

FORWARD: 5' GTAAGAGACGCGCAGT 3'  
BACKWARD: 5' CACACAGGAACAGCATATACCAT 3'.  
Location/Qualifiers  
1..474

## FEATURES

source

/organism="Antrodia cinnamomea"  
/mol\_type="mRNA"  
/strain="SMV 1"  
/db\_xref="taxon:279009"  
/sex="Heterokaryon"  
/cell\_type="Mycelium"  
/lab\_host="Cinnamomum kanehira"  
/clone\_lib="Antrodia cinnamomea cDNA library"  
/note="Vector: ZAP Express Vector"

## ORIGIN

Query Match 4.0%; Score 55; DB 8; Length 474;  
Best Local Similarity 54.1%; Pred. No. 0.00035;  
Matches 112; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1171 AATGCCGGGCGTATGATGATGATCGGCTGAGCTTAGAGTTCACTGCAAGCTGCC 1230  
DB 453 AATACCGGAGCTGGCGACATCATCTGCCCATCAGCATTCATTCAGACTGCC 394  
QY 1231 AAATGACAATGACGCGCTGTGGCAGTTTATCTGCCAAGTGAATGGATTACCAT 1290  
DB 393 CGTGTGCGCTTCAACGCTTCGGGTAAATTTATCAGCGCGGTCTCATAGCGTATC 334  
QY 1291 GATGATGCTTAATCTTGTGACCGTGAATGACCAAAAATTCACAGCTGACAGTCCAGACT 1350  
DB 333 CATTAACGCGCGCGCTGACACCGGAGCGAGAAAGTCCACGCAAGAAAGCCGGTGA 274  
QY 1351 GCCCAACGAGAGCGGCATATTTT 1377  
DB 273 GCCAGCGCATATATACACATTTT 247

RESULT 9  
B2559806/c 1171 bp DNA linear GSS 17-DEC-2002  
LOCUS pac82-164-1863.x1 pac82-164 Pseudomonas aeruginosa genomic clone  
DEFINITION pac82-164-1863, genomic survey sequence.

ACCESSION B2559806  
VERSION B2559806.1 GI:27176698  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonas aeruginosa

REFERENCE 1 (bases 1 to 1171)  
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.  
TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library J. Bacteriol. (2002) In press

JOURNAL Contact: Chris K. Raymond  
COMMENT Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

FEATURES  
source Location/Qualifiers  
1..1171

/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pac82-164-1863"  
/clone\_lib="pac82-164"  
/note="Clinical isolate 2-164 Whole genomic shotgun library."

## ORIGIN

Query Match 3.8%; Score 52.6; DB 9; Length 1171;  
Best Local Similarity 49.0%; Pred. No. 0.0022;  
Matches 170; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

177 CTTTGAGTAAATTTGCTCAATATCGGGCATCTTGGACGATTAATACACTATCA 236  
177 |||||  
396 CGTTGCCAGTGGCTGTCAATCAGCTAGGTGTACGACACACAGTACCGCCGCG 337  
237 CCCATGTCATATATGCTGACAGCTCACTAAATCTTACCAAGCCCA 236  
237 |||||  
336 AGAAGCGCGGACCGGCTCGGACGTCATGAAATGCGCCCTTGCCTTCGCGG 277  
336 |||||  
297 AATCAATGCGCACTGATATGATATGCAATTAACCGTGCACAGCCCATCAATTTGTCATG 356  
297 |||||  
276 CGATCAGACACAC---CTTGCCTGATGTCGAGACACAGCCCTGATGCGCGCAGG 220  
276 |||||  
357 TTGAGCGGAGTTGTTGATACCTTTGATGATCAATTAATATGCTGTATCAATATCA 416  
357 |||||  
219 CGGCGCGGAGCTTGTGGCTTGTGAATCGTCGTAGTACGACGCTCGCTCGCTA 160  
219 |||||  
417 TATATTCACAGCATGCGGTAGCCCTTTGAAGCTTAAAGGATACAGATGCGCATCA 476  
417 |||||  
159 CCCACTGCGCATGATGACCATGACGATCCGACAGACGCTTTCACGCGCGGATGCGTAA 100  
159 |||||  
477 TGGGAAGATTGGCAGCTGTGCTTAAGGCAATGCAATGAGGATTG 523  
477 |||||  
99 ACGGACGCGGCGGACCGATGCGCCAGTGCAGGCTCGGAGAGCGGTTG 53  
99 |||||

RESULT 10 838 bp DNA linear GSS 13-MAY-2005  
LOCUS CZ545153  
DEFINITION SRAA-aad60c01.g1 Strongyloides ratci whole genome shotgun library  
(SRAAGS 004) Strongyloides ratci genomic, genomic survey sequence.  
ACCESSION CZ545153  
VERSION CZ545153.1 GI:64677289  
KEYWORDS GSS.  
SOURCE Strongyloides ratci  
ORGANISM Strongyloides ratci  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Pangroloaimoidea; Strongyloidea; Strongyloides.  
1 (bases 1 to 838)  
Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Page,D.,  
Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,  
Clifton,S.W. and Wilson,R.  
Genome Survey sequences from the rat parasitic nematode  
Strongyloides ratci  
Unpublished (2005)

JOURNAL COMMENT  
Contact: Mitreva M  
Washington University in St. Louis  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: nematode@watson.wustl.edu  
Genomic DNA was provided by Fiona Thompson  
(F.Thompson@bristol.ac.uk) and Mark Viney  
(Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,  
UK.  
Classes: shotgun.  
Location/Qualifiers  
1..838  
/organism="Strongyloides ratci"  
/mol\_type="genomic DNA"  
/strain="isofemale line EBJ321 heterogonic"  
/db\_xref="taxon:34506"  
/dev\_stage="infective larval stage (L3)"  
/lab\_host="GS10"  
/clone\_lib="Strongyloides ratci whole genome shotgun  
library (SRAAGS 004)"  
/note="Vector: POTW13; Site\_1: BstXI; Site\_2: BstXI;  
Strongyloides ratci genomic DNA was randomly sheared,

end-repaired and size fractionated to enrich for 2-4 kb  
fragments. Genomic DNA was provided by Fiona Thompson  
(F.Thompson@bristol.ac.uk) and Mark Viney  
(Mark.Viney@bristol.ac.uk) at the University of Bristol,  
Bristol, UK. Sequencing by Washington University Genome  
Sequencing Center, St. Louis, MO."

ORIGIN  
Query Match 3.6%; Score 50.2; DB 10; Length 838;  
Best Local Similarity 62.2%; Pred. No. 0.01;  
Matches 79; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

946 GCCAATATTTACGCGCGACCAACAATATGCTGCATGTTGGCATCTCGCTACTAA 1005  
946 |||||  
76 GCCAATATTTCCGCTTAACCAACAATCAATCTTCATTTTCAGATGTGATGATTAA 135  
76 |||||  
1006 TGTCTGAACAGTGTCTTTGGCATTTGAGCTGTGATGCGGATGTTGCTGTTTTC 1065  
1006 |||||  
136 AAGCGTTCGTGTTTTCGTTGCTTCTGTGATCGCAATGATTTTCGATCGGTAA 195  
136 |||||  
1066 TTGAGTT 1072  
1066 |||||  
196 TTCAGAT 202  
196 |||||

RESULT 11 572 bp DNA linear GSS 09-APR-2002  
LOCUS BH795426  
DEFINITION EMBACJ78G09T7 PSU Brugia malayi Genomic Bac Library 3 Brugia malayi  
genomic, genomic survey sequence.  
ACCESSION BH795426  
VERSION BH795426.1 GI:20043758  
KEYWORDS GSS.  
SOURCE Brugia malayi  
ORGANISM Brugia malayi  
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
Onchocercidae; Brugia.  
1 (bases 1 to 572)  
Whitton,C., Daub,J., Quail,M., Hall,N., Foster,J., Ware,J.,  
Ganatra,M., Slatko,B., Barrell,B. and Blaxter,M.  
A genome sequence survey of the filarial nematode Brugia malayi:  
repeats, gene discovery, and comparative genomics  
Mol. Biochem. Parasitol. 137 (2), 215-227 (2004)  
15383292  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Main Road, Edinburgh, EH9  
3JT, UK  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk

Sequenced from the Brugia malayi BAC library constructed by Claire  
Whitton and Dr Mike Quail. The sequence was generated by The  
Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in  
collaboration with Mark Blaxter, ICAHB, University of Edinburgh,  
Edinburgh, UK  
Seq primer: T7 (TAATACGACTCATATAGCG)  
Classes: BAC ends.  
Location/Qualifiers  
1..572  
/organism="Brugia malayi"  
/mol\_type="genomic DNA"  
/strain="T8S"  
/db\_xref="taxon:6279"  
/sex="Mixed (male and female)"  
/tissue\_type="whole parasite"  
/dev\_stage="microfilaria (L3)"  
/clone\_lib="Brugia malayi Genomic Bac Library 3"  
/note="Vector: pBACe3.6; Site\_1: BamH I; Brugia malayi  
genomic DNA was partially cleaved with Sau3A I and size  
fractionated. 7,392 clones were generated with mean insert  
size ~48 kbp. The library was constructed by Claire

FEATURES  
source

1..838  
/organism="Strongyloides ratci"  
/mol\_type="genomic DNA"  
/strain="isofemale line EBJ321 heterogonic"  
/db\_xref="taxon:34506"  
/dev\_stage="infective larval stage (L3)"  
/lab\_host="GS10"  
/clone\_lib="Strongyloides ratci whole genome shotgun  
library (SRAAGS 004)"  
/note="Vector: POTW13; Site\_1: BstXI; Site\_2: BstXI;  
Strongyloides ratci genomic DNA was randomly sheared,



Whitton, Blaxter Nematode Genetics Lab, University of Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing Unit, The Sanger Centre, Cambridge, UK."

| Query Match           | 3.6%  | Score 49.8   | DB 9                                  | Length 572                       |
|-----------------------|-------|--|---------------------------------------|----------------------------------|
| Best Local Similarity | 58.4% | Pred. No. 0.012  |                                       |                                  |
| Matches               | 87    | Conservative   | 0                                     | Mismatches 62; Indels 0; Gaps 0; |
| QY                    | 924   | TTAATAATTCAAGTCAGCGCGTGC                                     | CAAAATTAATCCGCGCACCAACATCAATGACCTGCAT | 983                              |
|                       |       |  |                                       |                                  |
| DB                    | 206   | TTTCTAGATCCAAATATGAGAC                                       | CCTTAATTTCCACCAATAGCTATCTTTTCCCGGAG   | 265                              |
|                       |       |  |                                       |                                  |
| QY                    | 984   | GTTTTCGCATCTCGCCTTACTAATGTCGTAACAGTCCTTTGGCATTTGAGCCTGTGATGG |                                       | 1043                             |
|                       |       |  |                                       |                                  |
| DB                    | 266   | ATTTAAATATGTGACTTAATTAATGACGAGAGGTGATTTGCCAATCGTTCTGTGAGCG   |                                       | 325                              |
|                       |       |  |                                       |                                  |
| QY                    | 1044  | CGATGATTTGGTGTGTTGGCTGAGTT                                   |                                       | 1072                             |
|                       |       |  |                                       |                                  |
| DB                    | 326   | CTACAAATTTTCGTTCTTAGCTTAGCT                                  |                                       | 354                              |
|                       |       |  |                                       |                                  |

|            |                           |  |
|------------|---------------------------|--|
| RESULT     |                           | 12                                       |
| Locus      | BEI84674/c                |  |
| Definition | BEI84674                  | 452 bp mRNA linear EST 22-JUN-2000       |
| Accession  | PW4-HN0668-050500-002-d10 | HT0688 Homo sapiens cDNA, mRNA sequence. |
| Version    | BEI84674                  |  |
| Keywords   | BEI84674.1 GI:8663858     |  |
| Source     | EST.                      |  |
| Organism   | Homo sapiens (human)      |  |
|            | Homo sapiens              |  |
|            | Homo sapiens              |  |

|                   |  |
|-------------------|--|
| TITLE             | Shotgun sequencing of the human transcriptome with ORF expressed sequence tags |
| JOURNAL PUBLISHED | Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)                         |
| COMMENT           | 1073/7800<br>Contact: Simpson A.J.G.   |

Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?pt1&t2=PMW-HR0688-050>  
 500-002-d10ct3=2000-05-05&t4=1  
 Seq primer: puc 18 forward  
 High quality sequence start: 26  
 High quality sequence stop: 89.

```

FEATURES
source
    location/Qualifiers
    1..452
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /dev_stage="Adult"
    /clone_lib="H0688"
    /note="Organ: head; neck; Vector: puc18; Site_1: SmaI;
    Site_2: SmaI; A mini-library was made by cloning products
    derived from ORNSTES PCR (U.S. letters patent application
    No. 196,716 - Ludwig Institute for Cancer Research)
    profiles into the puc 18 vector. Reverse transcription of
    tissue mRNA and cDNA amplification were performed under

```

low stringency conditions."

|                       |        |   |        |                 |
|-----------------------|--------|---|--------|-----------------|
| Query Match           | 3.6%;  | Score 49.4;   | DB 2;  | Length 452;     |
| Best Local Similarity | 50.5%; | Pred. No. 0.014;  |        |                 |
| Matches               | 145;   | Conservative  | 0;     | Mismatches 141; |
|                       |        |   | Indels | 1;              |
|                       |        |   | Gaps   | 1;              |
| QY                    | 332    | CCAAAGCCATCAATTGCTGCAAGTGTGAGCCGATGTTTGACCTTTTGAAATCATT           | -AA    | 390             |
| Db                    | 435    | CCAGAGCCCTGCAGTSGCCCTCAAGAGCCGCGCACCGACGTTGGTGGCTTTGGACATCTGTTA   |        | 376             |
| QY                    | 391    | ATTATGCCCTTGATCAATATCATCATATATATTCACAGAGATGCGGTAGCCCTTTGAAGT      |        | 450             |
| Db                    | 375    | ATTACTGATCTTCATTGAGATGCGCGCACCCANTTGGCAGAGATGCTCAAGCCCGCATAGT     |        | 316             |
| QY                    | 451    | CTTAAAGGATCAAGCATGCGCATCATGCGGAAGATTGGCACTGTGCTTAAGCGAATGC        |        | 510             |
| Db                    | 315    | GCGCAGGCTCGCGAGCATGGCGGTGAATTTGGACAGCCAAACCGCGTGAACCGAATGCGCAAGGC |        | 256             |
| QY                    | 511    | AGATTAAGGCATTTGAGTAGTATGCTTTGACCTTTGATTTTAAATGATGATTTGGTAAAG      |        | 570             |
| Db                    | 255    | CGCCAGGCGCATTTGACATGCTGTTGTGTGGCGCACGAGATGTTCACTTCGGTACCGGATCGG   |        | 196             |
| QY                    | 571    | TTTGTTTTTTGGATGATATATATGCAATCAATCAATCAATCAATA                     |        | 617             |
| Db                    | 195    | GTTCCTGAATTCGTACGCGCAAGTACTTCTGCGCTTCTTTCACGTA                    |        | 149             |

|            |  |
|------------|--|
| RESULT     | 13   |
| CNS001A0/c |  |
| LOCUS      |  |
| DEFINITION | CNS001A00 1101 bp DNA linear GSS 03-JUN-1998                     |
|            | Drosophila melanogaster genome survey sequence TERN3 end of BAC: |
|            | BACR3D23 of RPL1-88 library from Drosophila melanogaster (fruit  |
|            | fly) genomic survey sequence.                                    |
| ACCESSION  | AJ068607   |
| VERSION    | AJ068607.1 GI:4958689  |
| KEYWORDS   | GSS.   |
| SOURCE     | Drosophila melanogaster (fruit fly)                              |
| ORGANISM   | Drosophila melanogaster  |

| REFERENCE           | AUTHORS | TITLE   | JOURNAL                                   | COMMENT |
|---------------------|---------|---|---|---------|
| 1 (bases 1 to 1101) |         | Genoscope.<br>Direct Submission                                     |   |         |
|                     |         | Submitted (02-JUN-1999)   | Genoscope - Centre National de Sequencage |         |
|                     |         | BP 191 91006 Evry Cedex - FRANCE                                    | (E-mail : segrete@genoscope.cns.fr        |         |
|                     |         | - Web : www.genoscope.cns.fr)                                       |   |         |
|                     |         | Determination of this BAC-end sequence was carried out as part of a |   |         |

```

FEATURES
    source
        location/Qualifiers
            1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR32D23"
            /clone_1lb="RPG1-98"
            /note="end : TGT3"

```

ORIGIN

```

Query Match      3.4%; Score 47.6; DB 10; Length 1101;
Best Local Similarity 18.8%; Pred. No. 0.061;
Matches 89; Conservative 178; Mismatches 205; Indels 1; Gaps 1;

Oy    398 TTGCATCAATTCATCATATATATTCACAGCGCATGCGGTAGCCCTTGAAGTCTTAAG 457
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    1081 KTKGTTKDDNKANVDAAKARDDDKTRAKTGMRKKKTTTTTKTKTKDMAAAAAAW 10222
       ::::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    458 GTATCAAGCATTGCATCCATCGAAGAATTGGCAGCTGTGCCTPAAGCAATGACAATAG 517
       ::::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    1021 AATKHKTDTTKGAKRTTKWGKTNWTKTKCKGTGKRGDAAKAAAADPKTDKRAKX 962
       ::::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    518 GCATTGAAGAGTTATNGCTTCCTTGATNTTTTAATTGATGATTGTAAGTTGTTT 577
       ::::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    961 KKKKKGRKKKKKKKKKKTKDGTKKKTKTGGKDDDDAMWWADTDWKAGGGGKGDA 902
       ::::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    578 TTTTGATGATATAAATGACCATACATCACCATCATCAATAATAAATAATGTCGGTTTG 637
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    901 WTDYTDMDAMWKGKQKOTKKKKKKGGGKAKRABAAAADRKAADDGPAAATKTKTKAK 842
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    638 TGCGTAAGTAAGCTATTTGTTGTGATVAAVTGTTGATTTGGGCAACGCTTGCAGTG 697
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    841 TWGKKKKGGGWGRKGTGRGKKGGGTGRAAGDTKKKKGGKKKKGTGDKRXTMWKTKTTG 782
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    658 CTCAGACATTCCTTGGCTAAATTCGCATCGTCTTGATGATGATGCGCGTTGGCAATTT 757
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    791 KTGTMTKTKTKGKKKKKKKKGGKGGDWTKGTMDTKTKTKGKTGSGGKKKTKYKGTJK 722
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    758 TCAAAATATACGCAATTTTGTGCCAATAPATCCACATGCGCATGTGCGCATCAAGATGA 817
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    721 T-WAAWAATAKTKTKTKTKGATKADTKTKKKKKDKKGTGKMKKKTKGAKAMWTGDR 663
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Oy    818 TCAGCGGAGAGATTTAAATTTGTTGCCCTTGAGCTGCGCAATTTGAGATATG 870
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    662 TKDKKKKAAWAMDITTDGKTAKKADPTADYGKGAATXAAXKAGMADADG 610
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 14
LOCUS      CNS002ZMK/C
DEFINITION Drosophila melanogaster genome survey sequence SPE end of BAC
            BACN02P17 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL097862
VERSION     AL097862.1 GI:5609473
KEYWORDS   GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscoidora;
            Dipteroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segreifgenoscope.cns.fr
            - web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CBPH (Centre
            d'etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBelOBAC11.

FEATURES             Location/Qualifiers
     source           1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN02P17"
                     /clone_1tb="DrosBAC"
                     /plasmid="pBelOBAC11"

```

| ORIGIN                | /note="end : SP6"   |
|-----------------------|---|
| Query Match           | 3.4%; Score 47.2; DB 10; Length 1101;                               |
| Best Local Similarity | 36.9%; Pred. No. 0.079;   |
| Matches               | 76; Conservative 42; Mismatches 88; Indels 0; Gaps 0                |
| QY                    | 651 TATTTGTGTGAAATCGTGTGATTTGGGGCAAGCGTTGTACAGTGTGTACAGATGCT 710    |
| DB                    | 885 TTTTWTWAGAAAAAGCCGGTGGRRKGGGGGGCAACGCTGGTTGGGGSGRSMATTTTK 826   |
| QY                    | 711 TGCGCTAAATTTGGCATGCTTTGTGTAGATGATGGCGGTTTGCAATTTTCAAAATACGA 770 |
| DB                    | 825 KTTTAKXGGGAAGATYKGGMAWTTTAKKGGTWTKKTTSTTDTWAAAATKTMCSG 766      |
| QY                    | 771 AATTTGTGCAATATATCACCATGCGATGTGCGCATCAAGATGATCAGCGGAGAGAT 830    |
| DB                    | 765 WTTTWTYTCYTTTWTWMCAPRAAGSCCTTCWCSTKSYMTWTGKTCRGVACATGTGK 706    |
| QY                    | 831 TTAATAATTTGTGCCCTTTGAGCTGCG 856                                 |
| DB                    | 705 TTRRGGSCGAGCCCTTTCKGTHCY 680                                    |

[illegible]

each tissue were then pooled. cDNA was prepared from 20 micrograms of mRNA according to the full-length cDNA library construction method described by Carninci P. et al. (2000). Genome Research 10(10):1617-1630 and directionally ligated into the pBluescript II SK (+) XR vector digested with XhoI (5' end) and BamHI (3'). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation."

# ORIGIN

Query Match 3.3%; Score 46; DB 8; Length 823;

Best Local Similarity 59.0%; Pred. No. 0.16; Mismatches 0; Gaps 0;

Matches 79; Conservative 0; Indels 55; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| Qy | 515 | AAGCATTGAGTAGTATGCTTATGCTTTGATTTTAATTGATGATGATGTAAGTTG  | 574 |
| Db | 204 | AGGGAATCGCTAGCTTATCTTGTATATGATTTAAATTAAGCATCGAATTAATGTG | 145 |
| Qy | 575 | TTTTTTGATGATATAATGCGATACCATCATCATCAATAATAAATCTGCCGT     | 634 |
| Db | 144 | TTTGAAGTATCAGCTATGCGAACCATGTGCGATGATACAGCTTTAAATTGGGCTT | 85  |
| Qy | 635 | TGCTGCTAAGTAA   | 648 |
| Db | 84  | TGTTGAGTATTA  | 71  |

Search completed: May 12, 2006, 11:35:43  
Job time : 6046 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 09:25:47 ; Search time 279 Seconds  
(without alignments)  
8792.240 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380  
Sequence: 1 taagcgtgacatgttaa.....acgacgcataatttttctga 1380

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/1/ina/1\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/H\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/PC/TUS\_COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/PP\_COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/RE\_COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length  | ID | Description       |
|------------|--------|-------------|---------|----|-------------------|
| 1          | 1380   | 100.0       | 96109   | 3  | Sequence 35, Appl |
| 2          | 1370.4 | 99.3        | 1410    | 3  | Sequence 73, App  |
| 3          | 226.2  | 16.4        | 1416    | 3  | Sequence 1282, Ap |
| 4          | 160.4  | 11.6        | 1830121 | 3  | Sequence 1, Appl  |
| 5          | 160.4  | 11.6        | 1830121 | 3  | Sequence 1, Appl  |
| 6          | 160.4  | 11.6        | 1830121 | 3  | Sequence 1, Appl  |
| 7          | 125.2  | 9.1         | 13323   | 3  | Sequence 2830, Ap |
| 8          | 117    | 8.5         | 1371    | 3  | Sequence 7702, Ap |
| 9          | 117    | 8.5         | 1401    | 3  | Sequence 7861, Ap |
| 10         | 117    | 8.5         | 1450    | 3  | Sequence 1, Appl  |
| 11         | 100.6  | 7.3         | 1416    | 3  | Sequence 5247, Ap |
| 12         | 84.8   | 6.1         | 640681  | 3  | Sequence 1, Appl  |
| 13         | 83     | 5.1         | 1368    | 3  | Sequence 1574, Ap |
| 14         | 69.8   | 5.1         | 567     | 3  | Sequence 7787, Ap |
| 15         | 67     | 4.9         | 2402    | 3  | Sequence 881, App |
| 16         | 61.4   | 4.4         | 1372    | 3  | Sequence 1, Appl  |
| 17         | 58.6   | 4.2         | 1242    | 3  | Sequence 86, Appl |
| 18         | 53.6   | 3.9         | 2423    | 3  | Sequence 86, Appl |
| 19         | 53.6   | 3.9         | 2423    | 3  | Sequence 86, Appl |
| 20         | 53.6   | 3.9         | 2423    | 3  | Sequence 86, Appl |
| 21         | 53.6   | 3.9         | 2423    | 3  | Sequence 86, Appl |
| 22         | 53.6   | 3.9         | 2423    | 3  | Sequence 86, Appl |
| 23         | 53.6   | 3.9         | 2423    | 3  | Sequence 86, Appl |
| 24         | 49.6   | 3.6         | 536     | 3  | Sequence 3, Appl  |

|   |    |      |     |       |   |                     |                   |
|---|----|------|-----|-------|---|---------------------|-------------------|
| C | 25 | 49.4 | 3.6 | 699   | 3 | US-09-134-000C-741  | Sequence 741, App |
|   | 26 | 48.6 | 3.5 | 1320  | 3 | US-08-956-171B-80   | Sequence 80, Appl |
|   | 27 | 48.6 | 3.5 | 1320  | 3 | US-08-781-986A-80   | Sequence 80, Appl |
|   | 28 | 48.2 | 3.5 | 1188  | 3 | US-09-710-279-655   | Sequence 655, App |
|   | 29 | 48.2 | 3.5 | 1188  | 3 | US-09-710-279-655   | Sequence 655, App |
|   | 30 | 47   | 3.4 | 1374  | 3 | US-09-134-001C-1001 | Sequence 1001, Ap |
|   | 31 | 45.6 | 3.3 | 1368  | 3 | US-08-936-165A-78   | Sequence 78, Appl |
|   | 32 | 42.6 | 3.1 | 10486 | 3 | US-09-902-540-9337  | Sequence 9337, Ap |
|   | 33 | 42.6 | 3.1 | 10486 | 3 | US-09-902-540-9337  | Sequence 9337, Ap |
|   | 34 | 42   | 3.0 | 1407  | 3 | US-09-530-838-1     | Sequence 997, App |
|   | 35 | 41.8 | 3.0 | 6693  | 3 | US-08-961-527-195   | Sequence 1, Appl  |
|   | 36 | 41   | 3.0 | 4010  | 3 | US-09-710-279-3540  | Sequence 3540, Ap |
|   | 37 | 40.8 | 3.0 | 1141  | 3 | US-09-806-708B-22   | Sequence 22, Appl |
|   | 38 | 40.8 | 3.0 | 1275  | 3 | US-10-113-113-3     | Sequence 3, Appl  |
|   | 39 | 40.2 | 2.9 | 1350  | 2 | US-08-665-435A-1    | Sequence 1, Appl  |
|   | 40 | 40.2 | 2.9 | 1350  | 2 | US-08-665-435A-3    | Sequence 3, Appl  |
|   | 41 | 40.2 | 2.9 | 1350  | 2 | US-08-843-309-3     | Sequence 3, Appl  |
|   | 42 | 40.2 | 2.9 | 1353  | 2 | US-08-843-309-1     | Sequence 1, Appl  |
|   | 43 | 40.2 | 2.9 | 1353  | 3 | US-09-583-110-197   | Sequence 197, App |
|   | 44 | 40.2 | 2.9 | 1383  | 3 | US-09-107-433-1773  | Sequence 1773, Ap |
|   | 45 | 39.4 | 2.9 | 3444  | 3 | US-09-710-279-4356  | Sequence 4356, Ap |

## ALIGNMENTS

RESULT 1  
US-09-596-002-35  
Application US/09596002  
Patent No. 6632636  
GENERAL INFORMATION:  
APPLICANT: Lagace, Robert, E.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAVELLA CATARRHALIS GENOME  
FILE REFERENCE: PM-0008-4 US  
CURRENT APPLICATION NUMBER: US/09/596,002  
CURRENT FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: 60/140,121  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PERL Program  
SEQ ID NO 35  
LENGTH: 96109  
TYPE: DNA  
ORGANISM: M. catarrhalis  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte template ID No. 6632636 35  
PUBLICATION INFORMATION:  
US-09-596-002-35

Query Match 100.0%; Score 1380; DB 3; Length 96109;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |       |  |       |
|----|-------|--|-------|
| QY | 1     | TAAGGCTGACCATTTTAAACAATTATCAGCTGCTATTATTAATTTTAAATGATC   | 60    |
| DB | 11357 | TAAGGCTGACCATTTTAAACAATTATCAGCTGCTATTATTAATTTTAAATGATC   | 11416 |
| QY | 61    | CATGCTAGACAGAGGAGGACAAAGACAGAGCCGAGGAGGCTTGACGCTGCCAA    | 120   |
| DB | 11417 | CATGCTAGACAGAGGAGGACAAAGACAGAGCCGAGGAGGCTTGACGCTGCCAA    | 11476 |
| QY | 121   | TACTGCCGATCAAGTGAACAACCTGATATATTTAGATGCAATGATGATCACTTT   | 180   |
| DB | 11477 | TACTGCCGATCAAGTGAACAACCTGATATATTTAGATGCAATGATGATCACTTT   | 11536 |
| QY | 181   | GAGTAATTTGCTCAATATTTGGGGGATCTTGACGATTAATATCACTATCCACCA   | 240   |
| DB | 11537 | GAGTAATTTGCTCAATATTTGGGGGATCTTGACGATTAATATCACTATCCACCA   | 11596 |
| QY | 241   | CTGCATTAATATGCTGACAGCTCACTAAATCTTGACCTTTACGAAGCCACCAAAAT | 300   |





```

Oy      721  GGATGCTCTGGTAGATGATGCGGGTTGGCAATTTTCAAAAATACCAATTTTGTGCG  780
Db      672  GGATGCTCTGGTAGATGATGCGGGTTGGCAATTTTCAAAAATACCAATTTTGTGCG  613
Oy      781  CAAATATCCACCATGCGATCGGCCGATCAAGATGATCAGCGGAGATTTAAATTGT  840
Db      612  CAAATATCCACCATGCGATCGGCCGATCAAGATGATCAGCGGAGATTTAAATTGT  553
Oy      841  TGCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAACTTGA CAGCTCAAGAC  900
Db      552  TGCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAACTTGA CAGCTCAAGAC  493
Oy      901  CACCAATTCATTTTGGATTCGTTAATTAATTCAGTCAGCGCGGCCAATTTACCGCC  960
Db      492  CACCAATTCATTTTGGATTCGTTAATTAATTCAGTCAGCGCGGCCAATTTACCGCC  433
Oy      961  GACACCAACATATGCTGCTCATGTTTGGCATCTCGCTCTAATGTCATAACAGTCT  1020
Db      432  GACACCAACATATGCTGCTCATGTTTGGCATCTCGCTCTAATGTCATAACAGTCT  373
Oy      1021  TTTGGCATTTGACCTCTGTATGCGCATGATTTGGTGTGTTTGTGCTTAGATTAGATCTCG  1080
Db      372  TTTGGCATTTGACCTCTGTATGCGCATGATTTGGTGTGTTTGTGCTTAGATTAGATCTCG  313
Oy      1081  TGCTTTGAGTGTATCATGAATGAACAATGAACATGCGCTGATCA CAGAAATGCCTTAGCCTTT  1140
Db      312  TGCTTTGAGTGTATCATGAATGAACAATGACGCTGATCA CAGAAATGCCTTAGCCTTT  253
Oy      1141  GCGCGGCAATATCTGCGCGCTTTGTGGATTATGCGGGGCTGATGATCTGATCGGC  1200
Db      252  GCGCGGCAATATCTGCGCGCTTTGTGGATTATGCGGGGCTGATGATCTGATCGGC  193
Oy      1201  TGAGCTTAGAGTTCACTGTCCCAAGCTGCCAAATATGACATGAA CCGCTGTGGGCACTTT  1260
Db      192  TGAGCTTAGAGTTCACTGTCCCAAGCTGCCAAATATGACATGAA CCGCTGTGGGCACTTT  133
Oy      1261  ATCTGCGCAAGTAGATTTGGGATTTACCATGATGATCTTAATTGTGACCGGTATGAC  1320
Db      132  ATCTGCGCAAGTAGATTTGGGATTTACCATGATGATCTTAATTGTGACCGGTATGAC  73
Oy      1321  CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGCGCATATTTTGTGA  1380
Db      72  CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGCGCATATTTTGTGA  13

RESULT 3
US-09-328-352-1282/c
; Sequence 1282, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1282
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1282

```

|         | Query Match | Similarity   | Score | DB 3        | Length |
|---------|-------------|--|-------|-------------|--------|
| Beet    | Local       | 52.1%  | 226.2 | 1416        |        |
| Matches | 718         | Conservative   | 0     | Mis-Matches | 59     |
|         |             |  |       | Indels      | 66     |
|         |             |  |       | Gaps        | 7      |
| Qy      | 2           | AAAGCGTAAACCATTTTAAACAATTACCTCGCTATATATTTTAAATGATGCC       | 61    |             |        |
| Db      | 1409        | AACGATATGACGACGCAACAACTGCTGACCAAGCTATTATTAACCTTTAAACATATCA | 1356  |             |        |
| Qy      | 62          | ATGCTAGACACAGCAGGTGACAAAAGCACAGCCCTTAGGGGTCTTGACTGCGTCCAT  | 121   |             |        |

|    |      |   |      |
|----|------|---|------|
| Db | 1349 | AAACTGACAAAGCTGGAT-----ACAAATACAAATCTTCAAGCTGTGTTCAAC         | 129  |
| Qy | 122  | ACTGCCGATCAAGTATACCACTGATATATTTAGTGCAGAAATGTCATCACTTGG        | 181  |
| Db | 1297 | GTTCACACAGCTCCGACAGCTCTTTAAAGCGTTGCTGCATGTAAATTTTACTTGCCT-    | 1239 |
| Qy | 182  | AGTAAATCTGCTCAATTAATGAGGGCAATCTGACGATTAATATACACTATCCACCAC     | 241  |
| Db | 1238 | TGAATGGCTTATTCATATGACAGGGCACTTACCAATCAATPACACAACTTTGGCATAT    | 1179 |
| Qy | 242  | TGCAATACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGGCCACCCAAATC    | 301  |
| Db | 1178 | TTTTGATAGAGAACGTAAAGAGAAAAATCTTGTCTCTTGGCCCTGTCCACTTAATA      | 1119 |
| Qy | 302  | AATGCCA-----GTATGATATGCCATMAACCGTCCAGCCCATCAATTGCTGCATG       | 355  |
| Db | 1118 | AGTCAACTTACCTTTTTCATCTTCAATGACAGCACTAAGCATCAATTGCTGCAAGT      | 1059 |
| Qy | 356  | GTTCAGCCGATGTTTGTACTTTTGAATCATTAATAATATGCTTGTCTATCAATATCA     | 415  |
| Db | 1058 | GTAGACCAACAAATTAAGTACTTTTAAAGTCAATTAATAACGACATGATGACGGTATTA   | 999  |
| Qy | 416  | ATATATTCACAGCGATGGGTAGCCCTTTGAAAGTCTTAAAGGATATCAAGATGGCATCC   | 475  |
| Db | 998  | ACGATCTACAGCGGTGCTCTAAGTCTTTAAATTTTAAAGTTTCAAGCATTTGACTCC     | 939  |
| Qy | 476  | ATGGGAAGATTGGACGTGTGCTTAAGGCAATGACAGATNAGGCAATTAAGGTATATGC    | 535  |
| Db | 938  | ATAGGTAAACAAATGCTTCCCAATGTATAACAGCTAAAGCATTAAGTATGTATGCC      | 879  |
| Qy | 536  | TTGCTTTGATTTTAAATGATGATTTGTATAAGTTTGTTTTGTATGATATATATGCC      | 595  |
| Db | 878  | ATACCTGGAATATTAATCTGACCTTTTATTAACCGCTGTAAACGACGACGAAGCA       | 819  |
| Qy | 596  | ATACCATCACATCAATAAATAAATAATCTGCGTTTGTGTGCTTAAGATACATAAT       | 655  |
| Db | 818  | AGCGTACCTGGCATCTCTTAAACGCCAATTAATGAT-----TAAATCCGTGCT         | 767  |
| Qy | 656  | GTGTGATTAAGTGTGATTTTGGGGCAACGCTTGTCAGTGTGTCAGACATTCCTGGCT     | 715  |
| Db | 766  | TTAAACCAAGCTTTGCATTGTGTCTATCTGSAACAAGTGACGGCTTA-----          | 716  |
| Qy | 716  | AAATTTGCATGCTTGTGTAGATATGAGCGGTTTGGCAATTTTCAAAAATACGAATTT     | 775  |
| Db | 715  | ----ACGCATCATCTCGTAAATTAACAATTTTACGCGCTTGAAATAATACATGTTTT     | 660  |
| Qy | 776  | TGTGCCAATATATCACACATGCCATGTGCGCGATCAAGATGATCAGCGAGAGATTTAA    | 835  |
| Db | 659  | GCTTATGATATACCCACACATATTTCAATGACGGCTTAAATGGTCTTCACTCATATTTAGA | 600  |
| Qy | 836  | ATTGTGCGCCCTTGAAGTGCACAAATTTGAGATATGCTCAAGCTGAAATCTTGACGTCA   | 895  |
| Db | 599  | ACCACTGTACTCAGCGTTTAAATGAGAGGTGTTCTTAATGAAAGCTTTGATTAATCTCA   | 540  |
| Qy | 896  | AGACCAACCAATTCATATTTTGATGTAAATTAATTCAGATGACAGCGTGCCAATATTA    | 955  |
| Db | 539  | AGAACGATTAATCTGCG---TTGATCTTAAATTAATCTTAACGCTGTGCACCAAGTTG    | 483  |
| Qy | 956  | CCGCGCACACCAATCATATGCTGTGATGTTTTCGATCTGCGCTACTAATGTGTATACA    | 1015 |
| Db | 482  | CCGCTACTGCAATTTTCTTACCTGCATCTTATGCCATTAAGGCCAATTAAGTGTATACA   | 423  |
| Qy | 1016 | GTGCTTTTGGCATTTGAGCTGTGATGGCGATATTTGGTGTGTTGTTGCTTGAGTTAGA    | 1075 |
| Db | 422  | GTACTTTTGTGATTTGAACTGTATATGCCAATGCGCACATCAGTGGC-----          | 373  |
| Qy | 1076 | TCTGTGCTTGTAGTATCAATGAACAATTTGAACATGCTGTATCAAGAAATGCTTGA      | 1135 |
| Db | 372  | -----ACGGCGCAATTAATGATATATGCGCCACACAGAAATATACCTTTA            | 330  |
| Qy | 1136 | GCTTTGGCGCGCAATATCTGCGCGTTCTTGGGTAAATGCGGGGCTGATGATGATCTGA    | 1195 |
| Db | 329  | GCAATATCTCTCGAATTTTCCGGTAATGTGTGACAGGCGTGGGCTTAAATAATCTCT     | 270  |

```

Oy      1196  TCGGGTAGCTTAAAGTTACCTGTCAGAGCTGCCAAATGACATGAAACCCCTGTGGGC 1255
Db      269   TCTGCTTTGTAATTAATTTCTTGATCAAGCTGACCAAAACTGGTTTAAACACGCGGA 210
Oy      1256  AGTTATCTGCCAAGTAGAGATTTGGGAATTAACCATGATGATGCTAACTTGTGACCGTGA 1315
Db      209   ATCTGATCGTGTCCGGAGAGGTGTGGGGGGGGAATTCGGTTACAGCAACTTGGTAGCTTGT 150
Oy      1316  TTGACCAAAAATTTACAGCTGTGACAGTCCAGAGCTGCCAAACGACGACGGCATAT 1372
Db      149   TCATGCAGAAATTTACAGCAGAAACACCTGATATTTCCAAAGCTGTACAACTTTT 93

RESULT 4
US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557, 884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match      11.6%; Score 160.4; DB 3; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 9.2e-37;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Oy      1    TAAAGCGGAACCATGTTAACAAATTTATCATCCTCGCTCATTTATATATTTAAATGATC 60
Db      1204439 TAAACATTGAGGTAAATATCGTAAATTTCTTCGCCGGCTTTTCAAAAGAAAGCAAACTGATC 12043880
Oy      61    CATGTACACACAAGCAGGTGTCAAAGACACAGCCCTAGCGGTGCTTTGACTGCTGCCAA 120
Db      1204379  GAGACTTGCACAAGCAGGCGCAATATATACCA-----TATCTCCGCTTTGCA 12043322
Oy      121   TATGCGCCGATCAAGTGTACCACTGTATATATTTAGATGCCAAAGTGCATCACTTT 180

```

|    |         |  |         |
|----|---------|--|---------|
| Dp | 1204331 | GGTTGGGGGTAAAAATCTATCGCTGTGTCATTGTATCGAAACA-----ATTAACCTTTG      | 1204278 |
| Qy | 181     | GAGTAAATCTTGCTCAATTAATCGGGGCACTTTGACCGATTAAATACACTATCCACCCA      | 240     |
| Dp | 1204277 | CGATGAAAAATTTTGGCAGACACGCCAGCACTTCGACCAAAACAAATTAATATGTGTGG      | 1204218 |
| Qy | 241     | CTGCATACATATGTGCTGACAGCTCACTTAAATCTTTGACTTTTACCAAGCCCAACCAAAAT   | 300     |
| Dp | 1204217 | TTGATTATATTAATTCAGCTAATTTCTGAAAATAATCAGCCCTTTTCGCTCTCCGCTACGAA   | 1204158 |
| Qy | 301     | CATGCCAGTGTATGATTCATTAACCCGTGCCAAGCCCATCAATTGCTGCATGTGTGA        | 360     |
| Dp | 1204157 | CAATATGCAATTTAACCTCTCAATATA-----AAGCCCAAGCCATGACGAACTGTACT       | 1204107 |
| Qy | 361     | GCCGATGTTTGAACCTTTTGAATCAATTAATATGCTGTCTATCAATATCATCAATATA       | 420     |
| Dp | 1204106 | CCCCAATTTTGTTGCTTTTAGGTCACTTATATCAACGAATGCCATTAGCTGATGCACTAA     | 1204047 |
| Qy | 421     | TTCAACAGCGATCCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGATGGCATCATGGG      | 480     |
| Dp | 1204046 | TTGAAAAAGATGATCTAACCTTTGAAATGACGAAGTGGCGGTAGCAATTTGAATCTAAATT    | 1203987 |
| Qy | 481     | AAGATGGACACTGTGCTTAAGGCAAGTGCAGATTAAGCAATTGATGATGATTCCTGCC       | 540     |
| Dp | 1203986 | AATACCTTATACCTTGTGCGCAATGCTGTGCTGCCAAATGTCTAATTAATTAATGCGAAC     | 1203927 |
| Qy | 541     | TTTGATTTTATATGATGATTTGGTAAAGTTTCTTTTATGATATATAATGCAATACC         | 600     |
| Dp | 1203926 | AACCAATGTACTCTTCCACAAAGGTAAATACACTTCACTTTTACATTAATTAATTCCTT      | 1203867 |
| Qy | 601     | ATCACCATCATCAATTAATTAATTAATATGTC- CGTTTGGTGTCTAAGTAATTTGTTG      | 659     |
| Dp | 1203866 | GCCATTTTCAAGTTTAAAGCCCAATATATCCGCACTATTTTCCGCAAAACAAACGGTATGTTT  | 1203807 |
| Qy | 660     | TGATTAATGTGTGATTTGGGGCAACGGTGCAGTGTGTCAAGCATTTGCTTGGCTAAT        | 719     |
| Dp | 1203806 | CGCTTGATTTTCGTTTTCCCAAAGTCAGCC-----                              | 1203775 |
| Qy | 720     | TGGCATGCTCTGGTGAATGATGAGCGGTTTGGCAATTTTCAAAATAACGAATTTTGTG       | 779     |
| Dp | 1203774 | ---TATCTTCAATGTTCACACACTACTTAGCATTAATGATTAATATCGTAATTTTGCTT      | 1203718 |
| Qy | 780     | CCAAATATTCACCAATGCAATCCGTGCGGATCAAAATATATCAGCGGAAGATTTTAAATTTG   | 839     |
| Dp | 1203717 | GCGCAATATCTTCTTAATATCAATTAAGGATCAATATATCTTCACTACGTCACGTTCAAGCAAG | 1203658 |
| Qy | 840     | TTGCCCTCTGAGCTGCCAAATTTAGATATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA    | 899     |
| Dp | 1203657 | TGCGTGGCGGAGCTTTTAAGCTAATTAAGTGTGCTCAAGCTGAAAACTGAAAGCTCTAGTA    | 1203598 |
| Qy | 900     | CCACCAATTCATATTTTGGATCGTTAATAATCAAGTGCAGGCGGTGCCAATATTACCGC      | 959     |
| Dp | 1203597 | CATAAAGTTCCAAATCTTCATTC---AACATGACAAAGGCGGAAATCCCAATATTTTCCGC    | 1203541 |
| Qy | 960     | CGACACCAACATATGCTGCTCATGTTTGGCATCTCGCCCTACATATGTGCTAACAGTGC      | 1019    |
| Dp | 1203540 | CCATACCAACTTTTACACACGACAGCTTTCCGCATTTTCAATACATTAAGTATTCGTTAC     | 1203481 |
| Qy | 1020    | TTTTGGCATTTTGAAGCTGTGATGCGAGATGTGTGTGTGTGTTGCTTGAATTGAATCTC      | 1079    |
| Dp | 1203480 | TTTTTAACATTTGAACGTGTATATCCCAAAATTTGCTTTGTGCTGCGCGGCA-----        | 1203429 |
| Qy | 1080    | GTCCTTTGAGTGTATCAATGAACAATTTGAACATGTGCTGATCAGAGAAATGCTTGAAGCTT   | 1139    |
| Dp | 1203428 | -----GAATATTCMAATATGCGGATTAATCTTCACTCCGCTTAA                     | 1203388 |
| Qy | 1140    | TGGCGGCGACATATCTGCGCTTCTTGGGTTAATATGCGGGGCTGATGATTCATTCGCG       | 1199    |
| Dp | 1203387 | GTCGGGTTTGAATTTCTGTGTGTTTTTACCGAAGCCCTCGGCTTAATTAACAATATATTCGC   | 1203328 |
| Qy | 1200    | CTGAGCTTTAAGATTCACTGTCCAAGCTGCCAAATATGACATGAACGCTGTGGGCACTTT     | 1259    |
| Dp | 1203327 | TTTCAAGTAAACATCTCTGATTTTAATCACTACAGTATTAAGGGGATATTTTGAAGGAAGTT   | 1203268 |

```

Oy      1260  TATCTGCCAAGGTAGAGATTGGGATTACATGCATGATCTTAACCTTTGTGACCGGTGATTGA 1319
          |||||
Db      1203267  TATCATATCAAGGTAGAGATTTTTCGAGTATCAATTCACAGAAATATAGCCCTGTGGGATA 1203208
          |||||

Oy      1320  CCAAAAATTCACAGCTGACGATCGACGCTGCCCAACGACGACGCGCATATTTT 1377
          |||||
Db      1203207  AGAGTATATCCACACAGAAAGACCTGTTTCCAGAGCCGATGATATATATTTT 1203150
          |||||

RESULT 5
US-09-643-990A-1/c
: Sequence 1, Application US/09643990A
: Patent No. 6528289
: GENERAL INFORMATION:
: APPLICANT: Robert D. Pleischmann
:           Mark D. Adams
:           Owen White
:           Hamilton O. Smith
:           J. Craig Venter
: TITLE OF INVENTION: The Nucleotide sequence of
:                   the Haemophilus influenzae Rd Genome, Fragments
:                   thereof, and Uses Thereof
: NUMBER OF SEQUENCES: 1
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville,
: STATE: MD
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3 1/2 inch diskette
: COMPUTER: Dell Pentium
: OPERATING SYSTEM: MS DOS v6.22
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/643,990A
: FILING DATE: 23-Aug-2000
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/487,429
: FILING DATE: 1995-06-07
: APPLICATION NUMBER: 08/426,787
: FILING DATE: 1995-04-21
: ATTORNEY/AGENT INFORMATION:
: NAME: Kenley K. Hoover
: REGISTRATION NUMBER: 40,302
: REFERENCE/DOCKET NUMBER: FBI86P1C1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 301-610-5790
: TELEFAX: 310-309-8439
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1830121 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      11.6%; Score 160.4; DB 3; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 9.2e-37;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Oy      1  TAAAGCGTGAACCATGTTAACAACTTATCACTGCTCATTTATATTTTAAATGATC 60
          |||||
Db      1204439  TAAACATTGAGCTTAATGCGGTAAATTCCTCGCCGCGCTTTCAAAAAGAACAAACTGATC 1204380
          |||||

61  CATGCTGACACAACGACAGGTGACAAAGACAGCCCTAGCGGTGCTTTGACTGCGTCGCA 120
          |||||
          GAGACTTGACAAAGCAGCGACCAATATATPACA-----TATCTCGCGCTTGCA 1204332

```

|    |         |   |         |
|----|---------|---|---------|
| QY | 121     | TACGCCGATCAAGTGCACCACTGATATTTTGGATGCCAAATGTCATCACTTT            | 180     |
| Db | 1204331 | TGTTGGGGTAAAAATCTATCGCTGTCTCATTTGATCGAACCA-----ATTA             | 1204278 |
| QY | 181     | GAGTAAATCTTGCTCAATATATCGGGGAGCTGTGACCGATTAATAACACTATCCACCA      | 240     |
| Db | 1204277 | CGATGAAAATTTTGCACAGCGCGGACGATCTCGACCAAAACATTAACAAATATATGTGTG    | 1204218 |
| QY | 241     | CTGCATTAACATATGCTGACAGCTCATTAATCTTGAACCTTTTACCAAGCCCAACCAAT     | 300     |
| Db | 1204217 | TTGATTAATTAATTCAGCTAATCTGAAAAATACGCCCCCTTTTCCGCTCCGCTAGCAA      | 1204158 |
| QY | 301     | CAATGCCGATGATGATATGCCATTAACCGTGCACAGGCCCATCAATTTGTCGCAATGTGTGA  | 360     |
| Db | 1204157 | CAATATCAATTTTACCTTCATATA-----AAGCCGACCAATGACCAACTGTACT          | 1204107 |
| QY | 361     | GCCGATGTTTGTACCTTTTGAATCATTTAAATATGCTTGTCTATCAATATCATCAATATA    | 420     |
| Db | 1204106 | CCCCACATTTTGTGCTTAAAGTATTAATCCACGAATGCCATTTAGCTTAAATGACATA      | 1204047 |
| QY | 421     | TTCAACGCGATGCGGTAGCCCTTTGAAAAGTCTTAAGGATTCAGACATGCGCATTCATGG    | 480     |
| Db | 1204046 | TTGAAAACGATGATCTTAACCTTTGAATGACGAAAGTGGCGGTGCAATTTGAATCTAAAT    | 1203987 |
| QY | 481     | AAGATGCGAGCTGTGCTTAAGGACGATGCAAGTAAAGCATTTAGTATAGTTATGTGCC      | 540     |
| Db | 1203986 | AATACCTATAGCTGTGCGCAATGCTGTGTGTCGCAAAATGTCATTAATTAATGCGCAC      | 1203927 |
| QY | 541     | TTGTATTTTAATTTGATGATTTGGTAAAGTTTGTTTTGTGATGATTAATAGCATACC       | 600     |
| Db | 1203926 | AACCAATGTAGCTTCTTCAAGGTAAATTAATCACTTATCTTTTACATTAATATGTGCTT     | 1203867 |
| QY | 601     | ATCACATTCATCAATTAATTAATTAATATGTC-CGTTTGGTCTTAAGTAACTATTTGTGG    | 659     |
| Db | 1203866 | GCCATTTTCAGTTTTHTGCCAAATATCGGACATATTTTCCGCAAAAGAAACGATATGTT     | 1203807 |
| QY | 660     | TGATTAATGTGTGATTTTGGGGCAACGCTTGTCAATGTGTCAAGCATTTGCTGGCTAAAT    | 719     |
| Db | 1203806 | CGCTTGATTTTTCGTTTTTCCCAAAAGTCACGC-----                          | 1203775 |
| QY | 720     | TGGCATGCTCTGGTAGATGATGGCGGTTTGGCAATTTTCAAAATACGCAATTTTGTG       | 779     |
| Db | 1203774 | ---TATCTCATTTGTCAACACACTATCTTTAGCATTTATATTAATGCGTAATTTTGCTT     | 1203718 |
| QY | 780     | CCAAATATATCACATGCACTGTCGCGATCAAGATGATCAGCGGAGAGATTTTAAATTTG     | 839     |
| Db | 1203717 | GCGCATATATCTTCAATATCCATATAGGATTCATATGATCTTCAATGACGTTCAAGACAG    | 1203658 |
| QY | 840     | TTGCCCCCTTGAAGCTGCAAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA  | 899     |
| Db | 1203657 | TGCGTGCAGCGACTTTTAACTATTAAGTTGTCTCAAGCTGAAAACTTAAGAGCTCTTAGTA   | 1203598 |
| QY | 900     | CCACCAATCCATTTTGGATCGTTAATTAATCAAGTGAAGCGCGCAATTTTACCCG         | 959     |
| Db | 1203597 | CATTAAGTTCAACATCTTCAATTC--AACATATCAAAAGGGGAAATCCCAATATTTTCCG    | 1203541 |
| QY | 960     | CGACACCAACAAATCATCTCGCTGCAGTGTTCATCTCGCTACTTAATGTCTGAACAGTGC    | 1019    |
| Db | 1203540 | CCATACCAACTTTCACACAGCAGCTTTGCCCATTTTCAATACTAAATGAATGATTAACGATAC | 1203481 |
| QY | 1020    | TTTTGGCATTTGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG  | 1079    |
| Db | 1203480 | TTTTTACATTTGAACCTGTATATCCCAATTTGGCTTTGTGCTGCGCGGCA-----         | 1203429 |
| QY | 1080    | GTGCTTTGAGTGTATCAATGAACATTTGAACATCGCTGATCAACAGAAATGCCCTTGAGCTT  | 1139    |
| Db | 1203428 | -----GAATATATTCATATATGCGCGATATCTTCACTCCGCTTTAA                  | 1203388 |
| QY | 1140    | TGCGCGGACAAATATCTCGCGGTTCTTGGGTTTAATGCGGGGCTGATGATGATTTGAATCG   | 1199    |
| Db | 1203387 | GTGCGGTTTGAATTTTCTGCGTTTATACCGCAAGCCCTGGGCTTAATTAACATATATATCGC  | 1203328 |
| QY | 1200    | CTGAGCTTAAAGTTCACTGTCCAGCTGCACAAATGAACATGAACCGCTGTGGGCAATT      | 1259    |

Db 1203327 TTTCAGTAACCAATTCCTGATTTAACTACAGTATAGAGGATATTTTGAGGAAGTT 1203268  
QY 1260 TATCTGCCAAGTGAAGTATGAGATTACATGATGATGATTAATTTGACCGGATGA 1319  
Db 1203267 TATCATACCAAGTAGGATTTTTCGATATCATACACAGATTTTACCTGTTGGGATA 1203208  
QY 1320 CCAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAACCGACGACGCAATATTTT 1377  
Db 1203207 AGAGATATATCCACCAAGAAAGACTGTTTTCGCAAGCCGATGATATATATTTT 1203150

RESULT 6  
US-10-158-865-1/c  
Sequence 1, Application US/10158865  
Patent No. 6846651  
GENERAL INFORMATION:  
APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag  
Patent No. 6846651  
FILE REFERENCE: PB16P2C1D1  
CURRENT APPLICATION NUMBER: US/10/158,865  
PRIOR APPLICATION NUMBER: US 09/557,884  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US 08/476,102  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/426,787  
PRIOR FILING DATE: 1995-04-21  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1830121  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4747)..(4747)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (9921)..(9921)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (40810)..(40810)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44416)..(44416)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:

NAME/KEY: misc\_feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44975)..(44975)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45593)..(45593)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45732)..(45732)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (47036)..(47036)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51334)..(51334)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51602)..(51602)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51786)..(51786)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51805)..(51805)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (55369)..(55369)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (65309)..(65309)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (65313)..(65313)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (80024)..(80024)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (100091)..(100091)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (102696)..(102696)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (105121)..(105121)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (107248)..(107248)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (117136)..(117136)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature



Db 1203428 -----GAATATATCAATATCGCGATTAATCTTCCATCCGCTTAA 1203388  
Qy 1140 TGGCGGACAAATCTGCGGTTCTTGGTTAATGCGGAGCTGATGATCTGATCG 1199  
Db 1203387 GTGCGGTTGAATTTCTGGTTTATCCGAGCCCTGGGCTTAATCAATATATCGC 1203328  
Qy 1200 CTGAGCTTAAGAGTTCACTCTCCAGCTGCCAAAATGACAATGACCCCTGGGCACTT 1259  
Db 1203327 TTTCAGATTAACATTCCTGATTTAACTACAGTAAAGAGGATATTTTGAAGAACTT 1203268  
Qy 1260 TATCTGCAAGAGTATGATGATTAACATGATGATGATTAATCTTGTGAACCGTATGA 1219  
Db 1203267 TATCAATACCAAGTAAAGATTTTTCGATATCAATCAACAAATATTAAGCTTGGGATA 1203208  
Qy 1320 CCAAAAATTCACAGCTGACGTCAGAGCTGCCCAACGACGACGCAATATTTT 1377  
Db 1203207 AGAGATATATCACACAAAGAAAGACCTGTTTGGCAACCCGATGATGATATTTT 1203150

RESULT 7  
US-09-543-681A-2830/c  
; Sequence 2830, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETTON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 2830  
; LENGTH: 1323  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-2830

Query Match 9.1%; Score 125.2; DB 3; Length 1323;  
Best Local Similarity 53.5%; Pred. No. 1.1e-27;  
Matches 352; Conservative 0; Mismatches 273; Indels 33; Gaps 3;  
Qy 723 CATGCTCTGGATGATGATGCGGTTTGGCAATTTTCAAAATATACGCAATTTTGTCCA 782  
Db 658 CATCATCTGGCTATATGACACAGGTTTACATTAATCAATATGATTAATTTAGCTTAC 599  
Qy 783 AATAATCCACCATGCC---ATCGCCGATCAAGATGATGAGGAGAGATTTAAATTTG 839  
Db 598 GGTATGTTTAACTTAACGATGATGATGATGATGATGATGATGATGATGATGATGATG 539  
Qy 840 TTGCCCCCTGAGCTGCAATTTGAGATATGCTCAAGCTCAAAATTTGACGCTCAAGCA 899  
Db 538 TGGCAGCCGCGCTGTTAAGCTATAGTGTCTCTAATTAAGCTCGATAGCTCTAAGA 479  
Qy 900 CCACCAATTCATATTTTGGATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 959  
Db 478 CATATATGATTAAGGCTG---GTTAATTAAGTTAAAGAGGATACCAATATTTGCGAC 422  
Qy 960 CAGACCAACCAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019  
Db 421 CAGACCAACCAAGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 362  
Qy 1020 TTTTGGCATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1079  
Db 361 TTTTGGCATTTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 312  
Qy 1080 GTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139  
Db 311 -----CAAAATATGCTCAATATACCGAACAATTTCCATCTCATGATCTC 269  
Qy 1140 TGGCGGACAAATATCTGCGGCTTCTTGGGTTAATGCGGCGCTGATGATGATGATGATG 1199

Db 268 GCCCTTCACATTAATCTGAGGCTGATTAACCAATATCCGACCTTGAGACAAATTAATCAG 209  
Qy 1200 CTGAGCTTAAGAGTTCACTCTCCAGCTGCCAAATGACAATGACGCTGTGGGCACTT 1259  
Db 208 CATCATTAAGCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 149  
Qy 1260 TATCTGCAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319  
Db 148 TATCTAAACCCGCGGATGCTGCGAGTATCTATCACTTAAGATGATTAACCTTGGCGGA 89  
Qy 1320 CCAAAAATTCACAGCTGACGTCAGAGCTGCCCAACGACGACGCAATATTTT 1377  
Db 88 TAAAAAGTCAACGACGATGACAGTAAATGCTATGCTCCAGATTAACAACTTTT 31

RESULT 8  
US-09-252-991A-7702/c  
; Sequence 7702, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7702  
; LENGTH: 1371  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7702

Query Match 8.5%; Score 117; DB 3; Length 1371;  
Best Local Similarity 48.4%; Pred. No. 3.5e-25;  
Matches 506; Conservative 0; Mismatches 500; Indels 39; Gaps 5;  
Qy 17 TTAACCAATCTATACCTGCTGCTATTAATTTTAAATTTGATGCTAGCAACAACA 76  
Db 1349 TTGGCAACAGCGGCTCGGCTTCTTGAAGTTCTTGAATGCTCAGGCTCGGCGAGGCC 1290  
Qy 77 GGTGACAAAGACAGCCGATGCGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 136  
Db 1289 GGTGACAAAGACAGCCGATGCGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1239  
Qy 137 GTACCAACTGATATATTTTGAATGCAATGATGATGATGATGATGATGATGATGATGATG 196  
Db 1238 ACTGCTTGTCCAGCGTTGCGACGCGCACCAAGGATACCGGCTGCGCGATGCTGGGA 1179  
Qy 197 ATTAATGCGGCTATTTGACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 256  
Db 1178 ATCAAGCCCGGCTCAAGGCAAGTACACCGCCGAGAAAGCGCGACCGGCTGCG 1119  
Qy 257 GACAGCTCACTAAATCTTGAACCTTACCAAGCCCAACCAATTAATGCTGATGAT 316  
Db 1118 CGGAGTCAATGAAATGCGGCTTGTGCGCTTGTGCGCGCGAGCAACAGCTTGGCG 1059  
Qy 317 ATGCAATTAACCGTGCACGACCATCAATGCTGCAATGATGATGATGATGATGATGATG 376  
Db 1058 TGATGCT---CGGCAACCGAGCCCTGATGATGATGATGATGATGATGATGATGATGATG 1002  
Qy 377 TTTGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 436  
Db 1001 TTGGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 942  
Qy 437 AGCCCTTAAGATCTTAAGGATCAAGATGCAATGCAATGCAATGCAATGCAATGCAATG 496  
Db 941 AGCCGGAAGAGCTTTCAGCGCGCTGAGCAATGCTGCAAGGACGAGCCGATG 882



```

OY      497 CCTAAGGCAAGGCGAGATTAAGCATTTAGTACGTAGTATAGCTTACCTTTGATTTTAAATTGA 556
Db      881 CCGAGGCGACAGGCGCGCCGAGCGCGTTGGAATAGTGTGGCGGCCACGGAATCTTCAGTTGC 822
OY      557 TGGATTGATTAAGATTTGTTTTTTTGTGATGATATTAATGATCCATACATCATCATCAAT 616
Db      821 CCAACCGGCGAGAGCTGTTCGAACTGGAAACGCGACACCTTGGCCCGCTTCTTCGATC 762
OY      617 AAATTAATAATCTGCCGTTTGGTGGCTTAAGTAAGTATTTGTTGTGATTAATGATGATTT 676
Db      761 AGGCCGAAAGC-----CTTGAAGTCCGCGCTTGTTCAGGCCGGA 724
OY      677 GGGGCAACGCTTGTCAGTGTGCTCAAGCATTCGCTGCTAAATTGSCCATCGTCTTGTAG 736
Db      723 CGACGACGACGCGCACGCTATCGGCGATTCAGCGGTGGGCA--GGGCATCGCGCGGATTC 666
OY      737 ATGATGCGCGTTTGGCAATTTTCAAAAATACGCAATTTTGTGCCAAATTAATCCACATG 796
Db      665 ACCACGACTGTGGCGGGCACCGCGGAAGATCCGGTCTTGGCCAGGTGATGACGCATG 606
OY      797 CCATGTGCCCATCAAGATGATCAGCGGAGAGATTTAAATTTGTTGCCCTTGAAGCTGCC 856
Db      605 CCGTGTAGCGATTCATATGATGCTTCGTGTAAGTTCACACGCTGCCACCTCGCGCGTTG 546
OY      857 AAATTGAGATATATGCTCAAGCTGAAACCTTGACAGCTCAAGCAGCAACCAATTCATATT 916
Db      545 AGCGCATGCGCAGGTTTCCAGCTGGAAGCTCGACAGCTCCACACGTAAGCTCGATGTGC 486
OY      917 TGGATCGTTAATATTAATTCAGATGACGCGCTGCCAATATTACGCCGACACCAACATCATG 976
Db      485 TCG---GCCAGACGGTTCAGCGCGCGGGGTGCCAGAGTTGCCGCCGACGCGACGCTTG 429
OY      977 CCGCATGTTTTCCATCTCCGCTTACTATGTCGTAACAGTGTCTTTGGCATTTGAGCCT 1036
Db      428 TCCGCGGCCACCGCCATTTCCGCCACACGAGGTGTGACAGTGTCTTTCGCTTGGAAACCG 369
OY      1037 GTGATGCGCATGATGATGTGTGTTG 1061
Db      368 GTGATGCGCATGATGCGGGCCCTTCG 344

RESULT 9
US-09-252-991A-7861
; Sequence 7861, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7861
; LENGTH: 1401
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7861

```

|                       |       |  |        |                 |
|-----------------------|-------|--|--------|-----------------|
| Query Match           | 8.5%  | Score 117;   | DB 3;  | Length 1401;    |
| Best Local Similarity | 48.4% | Pred. No. 3,6-25;  |        |                 |
| Matches               | 506;  | Conservative   | 0;     | Mismatches 500; |
|                       |       |  | Indels | 39;             |
|                       |       |  | Gaps   | 5               |
| QY                    | 17    | TTAACAACCTATACACCTGGCATTTATATATTTAAATGTATGCATGTACACAGCA          | 76     |                 |
| DB                    | 14    | TTGGCAAAAGCGCGTCCGGGTTCTTGAAGTCTTGAACATGTCCAGGCTCCGCAAGCC        | 73     |                 |
| QY                    | 77    | GGTGCAAAAAGCAGACCCCTGAAGGGTGTCTTTGACTGCGTCGCCAATACTGCCCGCATCAAGT | 136    |                 |

|    |      |   |      |
|----|------|---|------|
| Db | 74   | GGCGACAAACGACCGGATCGCCTTGGC-----GGGCGAGTCGGCGGCGCTGGCG          | 124  |
| Qy | 137  | GTACCAACCTGATATATTTTGAATGCCAAATGTGCATACCTTTGAGTAAATCTTTGCTCA    | 196  |
| Db | 125  | ACTGTTTCGTCCAGCGTTGGCCAGCGGCAACGCGGTACCGGTTGGCCAGTGGCCTGGGCA    | 184  |
| Qy | 197  | ATTATCGGGGCGATCTTGCACGATTAAATPACACATATCACCCACTGCGATPAACATAGCT   | 256  |
| Db | 185  | ATCAGCCCCGGGCTCACGGCCAGACGATACACCGCCGGCAGAAACGCGCGACCGGCTCG     | 244  |
| Qy | 257  | GACAGCTCACTAAATCTTGAACCTTTACGAGCCCAACCAAAATCAATGCAAGTATGAT      | 316  |
| Db | 245  | CGCAGGTCATGAAATCGGCGCCCTTGGCGTCTCCGCGCGAGACGACCAAGCTTGGCG       | 304  |
| Qy | 317  | ATGCGATPAACCGTGGCCAAAGCCCATATTCATTCGTCGAATGTGTGAGCGCATTTGTACT   | 376  |
| Db | 305  | TCGATGT---CGGACCCAGCCCTTCGATGCGCGCAGGAGGCGGCGCAGCTTGGTGGCC      | 361  |
| Qy | 377  | TTTGAATCATTTAAATATATGCTTGTCTATCATATATCATATATATTCACAGCGATGCGGT   | 436  |
| Db | 362  | TTGGATGTGTGTGTAGCTCAGCGCCGTGCGCTGGGTACCACTGGCAACGATGAGCC        | 421  |
| Qy | 437  | AGCCCTTTGAAGTCTTAAAGGTATCAAGCATGGCATTCATGGGAAATTTGGCAGCTGTG     | 496  |
| Db | 422  | AGGCGGAGAAACGGCTTTCAGGCGCCGACGATGCGTGTGAACGGCAGGCCACCGCATGG     | 481  |
| Qy | 497  | CCTTAAGCGAATGCGAGATTAAGGCAATGATGATGATATGCTTGGCTTTGATTTTAAATGA   | 556  |
| Db | 482  | CCAGCGCCGACGGCGGCGAGCGCTTGGAAATGTTGTGGGCGCACGATCTTCAGTTGC       | 541  |
| Qy | 557  | TGATTTGGTAAAGTTTGTTTTGTGTGATATATAAGCATATACATCATCAATCAAT         | 616  |
| Db | 542  | CCAACCGGAGCACTTGTGCACTGGAACGCGACCACTTGGCGGCTTCTTCGATC           | 601  |
| Qy | 617  | AAATTAATATCTGCCGTTTGTGTGCTAATGATATTTGTGTGATATATGTGTGATTT        | 676  |
| Db | 602  | AGCGCCGAAGC-----CTTAAGTCCGGCTTGTCAAGCCCGAA                      | 639  |
| Qy | 677  | GGGCGAAACGCTGTTCAGTGTGTGTCAGACATTTGCTGGCTAAATTTGGCATCGTTTGGTATG | 736  |
| Db | 640  | CGACACGAGCGGACGGTATCGGCGATCAAGCGGTGGGTCA--GGGATCGGCGCGATTC      | 697  |
| Qy | 737  | ATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTGTGCGCAATTAATCAACATG     | 796  |
| Db | 698  | ACCAACGACTGGCGGGGACCGCGGAAATCCGGTGCCTTGGCAGAGTGTATGTCAGCATG     | 757  |
| Qy | 797  | CCATCGAGCGCATCAAGATGATCAGCGGAGAGATTTTAAATTTGTGCCCCCTGAGCTGCC    | 856  |
| Db | 758  | CCGTCGTAGCGATTCATATGTGTTCTTCGTGACGTTCAAGCAGGTGCGCACTCGGCGTTG    | 817  |
| Qy | 857  | AAATTTGATGATGTCTCAAGCTGAAAACCTTGACAGCTCAAGCACACCAATTCATATTT     | 916  |
| Db | 818  | AGCGGATCGAGGTTTCCAGCTGGAACTCGACAGCTCCAAACGTTACAGCTCGATGTGCG     | 877  |
| Qy | 917  | TGATTCGTTATATATTCAGTGCAGAGGGGTGCGAATATTACCGCGGACCAACAAATATATG   | 976  |
| Db | 878  | TCG---GCCAGAGGTGAGAGCGCGGGGTGCGAGTTTGC CGCGGACGCGACAGCTTG       | 934  |
| Qy | 977  | CTGCGATGTTTGGCATCTCGCTACTATATGTGTGTATCAGTGTCTTTTGGCATTTGAGCCT   | 1036 |
| Db | 935  | TTCGCGGCGCAACCGCATTTTCGCCACACAGGAGTGTGTCAGGTGTCTCTTCGGGTTGAACCG | 994  |
| Qy | 1037 | GTGATGGCGATGATTTGGTGTGTGTTGG                                    | 1061 |
| Db | 995  | GTGATGGCGAGATCGGGGCTTTCG  | 1019 |

RESULT 10  
US-09-701-229-1/c  
; Sequence 1, Application US/09701229  
; Patent No. 6890910  
; GENERAL INFORMATION:



```
APPLICANT: El-Sherbelini, Mohammed
APPLICANT: Azolina, Barbara
TITLE OF INVENTION: MURD PROTEIN AND GENE OF PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA
FILE REFERENCE: 20193P
CURRENT APPLICATION NUMBER: US/09/701,229
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 60/087,308
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: PCT/US99/11585
PRIOR FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PseBSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1450
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-701-229-1
```

```
Query Match      8.5%; Score 117; DB 3; Length 1450;
Best Local Similarity 48.4%; Pred. No. 3,6e-25;
Matches 506; Conservative 0; Mismatches 500; Indels 39; Gaps 5;
```

```
QY 17 TTACCAACTTATCCTGCTCATTAATTTTAAATGATCCATGCAAGCA 76
DB 1375 TTGGCCAAAGGCGTCGCGCTTTCGAAAGTCTTGAAACATGTCAGGCTCGCGCAGGCC 1316
QY 77 GGTGACAAAGACAGAGCCCTAGCGGTGCTTGACTGCGTCCAAATCTGCCGATCAAGT 136
DB 1315 GGGGACAAACAGACCCGATGCGCTTGGC-----GGGCCAGCTCGGCGCTGCGCG 1265
QY 137 GTACCAACTGATATATTTAGATGCCAAATGTGATCATCCTTTGAGTAATCTTGCTCA 196
DB 1264 ACTGCTTCGTCAGCGCTTGCGAGCGGCAACAGCGGATACGCGTTGCCAGTGTCTGGGCA 1205
QY 197 ATTATGGGGCATCTTGACCGATTAATACACTATCCACCACTGCAATACATATGCT 256
DB 1204 ATCAGCCCGCGTCAGCGCCAAACAGTACCAACCGCGGAGAGCGCGGACCGGCTCG 1145
QY 257 GACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAATCAATGACGATGATGAT 316
DB 1144 CGAGGTCATGAAATCGGCGCCCTTGCCGTCTCCGCGGAGAGACACCACTTGCCG 1085
QY 317 ATGCCATAAACCGTGCACAGCCCATCAATTCTGCAATGTGAGCGGATGTTGTACT 376
DB 1084 TCGATGT---CGGACCCAGCCCTCGATGCGCGGAGGCGCGCGGACGTTGTGGCC 1028
QY 377 TTGATCATTAATAATATGCTTGTCTATCATATCATATATTTACAGCGATCGCT 436
DB 1027 TTGAAATGTCGTAGTAGTCAAGCCCTGCGCTCGCTACCACTGCGAGCATGAGCC 968
QY 437 AGCCCTTGAAGTCTTAAAGGATATCAAGCATGAGCATCCATGGGAAGATGGAGCTGTG 496
DB 967 AGGCGCGAAACCGCTTTCAGCGCGCGGAGCATGCGCTTCAGACGCGGACCGCATG 908
QY 497 CCTAAGCAGATGACAGATTAAGGATTAAGGATTAAGCTTTCGCTTGAATTTTAAATGA 556
DB 907 CCAGAGCGCAGCGCGCGGAGCGGCTTGAATTAAGTGGGCGGCAAGATCTTCAAGTTC 848
QY 557 TGGATGTGTAAGATTTGTTTTTTTGAATATTAAGCCATACCATCATCATCAAT 616
DB 847 CCAACCGGAGCGAGCTTGTGCAACTGGAACGCCAGCATCTTGGCGGCTTCTCGATG 788
QY 617 AATAAATAATCTGCGCTTGTGGCTAAGTAAGCTATTTGTGTAATATGATGATTT 676
DB 787 AGGCGCAAAAGC-----CTTGAAGTCCGGCTTGTTCACGCCGAA 750
QY 677 GGGGACAGCGCTGTCAATGTGATCAAGCATGCTTGGCTAATATGCAATGCTTGTGATG 736
DB 749 CCAACGACAGCGGACGATATCGGCGATCAGCGGATCGGATCA--GGGATGAGGCGCATTC 692
QY 737 ATGATGCGGCTTGGCAATTTTCAAAATACGAATTTTGTGCGCAATATCATCAATG 796
```

```
DB 691 ACCACGACCTGGCGGCGCACCGGGAAGATCCGGTCTTGCGCAGTGTAGTCAGCCATG 632
QY 797 CCATGTGCGCATCAATGATGATACAGGAGAGATTTAAATTTGTCCTTGAGCTGCC 856
DB 631 CGGTGTAAGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
QY 857 AATTTGAGATATGCTCAAGCTGMAAACTTGACAGCTCAAGACCAACCAATCATATTT 916
DB 571 AGGCGATCGAAGGTTTTCAGCTGGAAGTCTGACAGCTCCAAACGATACAGCTGATGTCG 512
QY 917 TGGATGTTAATTAATCAAGTGAAGCGGTGCAATATTAACCGCGGACCAACATATG 976
DB 511 TCG--GCCAGAGGTGAGCGCGCGGCTGCGAGGTTGCCGCCAGCGGACGCTTG 455
QY 977 CTGCAATGTTTTCATCTGCGCTTCAATATGTCGTAACAGTGTCTTGCGATTTGAGCCT 1036
DB 454 TCCGCGGCGCACCGCATTTTCCGACCAAGGATGTACAGGTGCTTCCGCTTGGAACCG 395
QY 1037 GTGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1061
DB 394 GTGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 370
```

```
RESULT 11
US-09-489-039A-5247/c
Sequence 5247, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NOLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 5247
LENGTH: 1416
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5247
```

```
Query Match      7.3%; Score 100.6; DB 3; Length 1416;
Best Local Similarity 51.1%; Pred. No. 3,9e-20;
Matches 337; Conservative 0; Mismatches 289; Indels 33; Gaps 3;
```

```
QY 722 GCATGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
DB 746 GCATGCTCGCGCTTCAACGAGGATCTTCGATTCCTGTAATTCGCGAGCTTCCGCGG 687
QY 782 AATATATCACCATGCG---ATCGCGCATCAAGATGATGATGATGATGATGATGATGAT 838
DB 686 CGATACGCTGCGACCGCCCGAGGATGATGATGATGATGATGATGATGATGATGATG 627
QY 839 GTTGCCCTTGAAGCTGCAATTTGATATGATGATGATGATGATGATGATGATGATGATG 898
DB 626 GTTGCGGCAACCGCTGAGCGGTGAGGATGATGATGATGATGATGATGATGATGATG 567
QY 899 ACCACCAATTCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
DB 566 ACATACAGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 510
QY 959 CCGACACCAACATCATGCTGCAATTTTGCATCTGCGCTTCAATGATGATGATGATGATG 1018
DB 509 CCAAGCGGACGTTAATCTCAGGCGCTTTCGCAATCTCCCAACAGGATGATGATGATG 450
QY 1019 CTTTGGCATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1078
DB 449 CTTTATCGTTTGAACCGGATATGATGATGATGATGATGATGATGATGATGATGATG 400
QY 1079 CGTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
```

Db 399 -----GCAGACAGTTGCATATCGCCACGATCTTACCGCAGCTCG 357  
Qy 1139 TTGGCGGCGACATACACTCGGCTTCTTGCTTAATGCCGGGCTGATGATGATGATG 1198  
Db 356 GCCCTGCGCTGACGAGATGAGTGGCCAGCCGCAATTCGCCGATGGCAGATCAGGTC 297  
Qy 1199 GCTGAGCTTAAAGATTCACTGTCCAGCTGCGCAAAATGACAAATGAAAGCCCTGTGGGAGT 1258  
Db 296 GCCCGCAGCAGCCAGATATCAATGATGTCGCCGACATGCGCATGACCGATTCGGGTAT 237  
Qy 1259 TTATCTCCGCAAGTATGATTTGGGATTAACATGATGATGATGATGATGATGATG 1318  
Db 236 TTATTCAGCCCCGGAGGCGCGAGCGGTGTTCATCAGCGCGGGTTACGCCCGCGCC 177  
Qy 1319 ACCAAAAATTTCAAGCTGACAGCTCCAGAGCTGCCAAACGAGCGCATATTTTTT 1377  
Db 176 ATGAAAAAGTCCAGCAGTACAGGCGGTCTATGCCCGCCAAATTAAGACATTTTTT 118

## RESULT 12

US-09-790-988-1  
; Sequence 1, Application US/09790988  
; Patent No. 6632835  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: SHIGENOBU, HIDEKI  
; APPLICANT: HATORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 6.1%; Score 84.8; DB 3; Length 640681;  
Best Local Similarity 50.2%; Pred. No. 8e-14;  
Matches 289; Conservative 0; Mismatches 257; Indels 30; Gaps 2;

Qy 802 GTGCCCATCAAGATGATCAGGAGAGATTAAATTTGTCCTGAGCTGCCAAAT 861  
Db 240727 GTATCGATGATATGATCTTCGCTAAATTAAGAAATACGCTATCTTAATTTAAT 240786  
Qy 862 TGAGATATGCTCAAGCTGAAACTTGACGCTCAAGCAGCAACCAATTCATTTTGAT 921  
Db 240787 AAATGATTTTCTAGTTGAAAACAGACAGTTCTATATGATTAATCTGCTCTT--T 240843  
Qy 922 CGTTAATATTCAGAGTCAAGCGCTGCCAATATTCAGCAGCAGCAATCATGCTGCG 981  
Db 240844 ATCAAGATTTCTAGTACGGAACACCTATATTCACACCTAGAAAACCTTATATCTG 240903  
Qy 982 ATGTTTGCATCTCGCTTAATATGTCGTAACAGTCTTTTGCAATTTGAGCTGTGAT 1041  
Db 240904 TTTTGTGCAATTTTTCATATCATGTAAGTACAGTCTTTTTCATTAAGTCCCTGTA 240963  
Qy 1042 GCGGATGATGAGTGTGTTTGTGCTTGAATATCTGCTGCTTGAAGTATCAATGA 1101  
Db 240964 TGAATATATGAGACAGTCACTTCTTAG-----AAAA 240996  
Qy 1102 CAATTGAACATCGCTGATCAAGGAATCGCTTGAAGCTTGGCGGCGCAATACTCGGCGT 1161  
Db 240997 TAGTTCAATATCACTAATATTCATACCTAGCAACGCTTTTATTAATTAAGGTTT 241056  
Qy 1162 TCTTGGTTAATGCGGGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1221  
Db 241057 AAATGAAGAAATCCCGAGCTTATTAACAATTAAGTCTGAATTCAGATTCACATGATGTC 241116

Qy 1222 CAAGCTGCAAAATGACATGAACCGCTGCGGAGTTTATCTGCAAGGTAGATTGGG 1281  
Db 241117 TAACTCTTAACTATATTCATATTTTGAAGAAATTTAATTAATTTGAAGATGTTT 241176  
Qy 1282 ATTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341  
Db 241177 AGATTCAATATATTTTAAAGTTTATTTTAAATTTTAAATTTAATGCAAGATAT 241236  
Qy 1342 TCCAGAGCTGCGCAACCGAGCAGCGATATTTTTT 1377  
Db 241237 TCTGTAAACCATCTCTAAATTAATTAATTTTTT 241272

## RESULT 13

US-09-107-532A-1574/C  
; Sequence 1574, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; ADDRESS/SEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 1574:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1368 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (b) LOCATION 1...1368  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1574:  
US-09-107-532A-1574

Query Match 6.0%; Score 83; DB 3; Length 1368;  
Best Local Similarity 53.6%; Pred. No. 9.5e-15;  
Matches 173; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy 312 ATGATATGCAATTAACCGTGCAGCCCATCAATTCCTGCAATGTTGAGCCGATGTTG 371

```
Db      1069 ATAAACGACGCTTAGACGTTTCAATCCACTTAGAGCCATTGTCGCTAAATATGCG 1010
Qy      372 TACCTTTGATCATTAATAATATGCTTGCTATCAATATGATCATTAATATGACAGCAT 431
Db      1009 TTGCTTTGAATATATTAATAATTTTCTGCTTGATTTCTTCTACATATTTGGGTACGAT 950
Qy      432 GCGGAGCCCTTTGAAAGTCTTAGGGATCAAGCATGGCATCCATGGGAAGTTGGCAG 491
Db      949 GTGGTAGCCCATGGAATAATGCAACGTTTCTGTAAGCTTCATTTGATATACATACA 890
Qy      492 CTGTGCTTAAGGAAGTGCAGATGAGCATGAGTAGTATGCTTACCTTGTATTTTAA 551
Db      889 ATTTTGTCTCAGAAATAGCAGCTAGTGCATTTTCTACATTATGCTTACCTGTAACCCA 830
Qy      552 ATGATGATGTTGGTAAAGTTGTTTGTGATGATATTAATGCGATACATCATCAT 611
Db      829 GTTCAATATTTTCCATTTATTTTCTTTTATAGTATATAGTAAATGCTCAGCATACA 770
Qy      612 CAAATTAATTAATAATCTGCGTT 634
Db      769 CACCGTCTTCAACACTTCTTT 747

RESULT 14
US-09-252-991A-7787
; Sequence 7787, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7787
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7787

Query Match      5.1%; Score 69.8; DB 3; Length 567;
Best Local Similarity 55.0%; Pred. No. 6.5e-11;
Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy      327 CCGTGCACAGCCCATCATTTGCTGCAATGGTTGAGCCGATGTTGTACTTTTGAATCAT 386
Db      262 CCGCACCCAGCCCTGATCGCCGACGCGCGCGAGCGTGGTGGCTTTGAAATCG 321
Qy      387 TAAATATATGCTTGTATCAATATCATCAATATATTCACAGCATGCGGTAGCCCTTTGA 446
Db      322 CGTAGTAGTCAAGCCCTCGCTCGCTGTAACCATGCGAGCATAGCCAGGCGCGAATA 381
Qy      447 AAGTCTTAAGGATGATCAAGCATGAGCATGGAATGAGGATGAGCGTGGCTTAAGGCA 506
Db      382 AGGCTTTCAGCGCGCGAGCATGAGCTTCAACGCGAGCGCATCGATGCGCCAGCGCA 441
Qy      507 GTGCAATTAAGGATGAGTATGCTTGTGCTTGTGATTTTAAATGATGATGTTGA 566
Db      442 GCGCGGCGAGCGGCTTGAATAGTATGTTGGGCGCCAGGATCTTCAGTTGCGCAACGGCA 501
Qy      567 AAGTTTGT 575
Db      502 GCGAGTTGT 510

RESULT 15
US-09-221-017B-881/C
; Sequence 881, Application US/09221017B
```

```
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 881:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...2402
US-09-221-017B-881

Query Match      4.9%; Score 67; DB 3; Length 2402;
Best Local Similarity 46.6%; Pred. No. 1e-09;
Matches 334; Conservative 0; Mismatches 365; Indels 18; Gaps 3;

Qy      355 GGTGAGCCGAGTGTGTTGATCTTTGAATCATTAATAATGCTTGTCTATCATATCATC 414
Db      2165 GGTAGAAATTTAGTGTGTGATGTTGAGTCAATGATATATGACACCCCTTTACGGGAC 2106
Qy      415 AATATATTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGATC 474
Db      2105 AATCTTTTCAGTGCATGCGGTATCATTTCTTGAATCTTCAAGGCTTCTCGAATAGCCTC 2046
Qy      475 CATGGAGATTGGCAGCTGTGCTTAAGGCAAGTGCAAGTGAAGGATTAAGTATGTTATG 534
Db      2045 ATTTCTTAATATCAATAGCTTTGGCAGCAATGCGAGTGGCATAGCATGTGCTGTTGTG 1986
```

QY 535 CTGCGCTTGAATTTTAAATGATGATGTAAAGTTTGTTTTGTGATATATAATGC 594  
 Db 1985 CATGCCGAGACAGAGCTAAAGTTCTCATCATTAACAAGGTGAATTAAGATTCATGAC 1926  
 QY 595 CATACCATCACCATCATCAATTAATTAATAATCTGCCGTTGTGGCTAAGTAAGCTATT 654  
 Db 1925 AACCAATTGCTTCTTTTCATTGATCAAGCAGTTGTATATCCGTCGAGCTCCATGTC 1866  
 QY 655 TGTGTGATATAGTGTGATTTTGGGGCAACGCTTGCAGTGTGTCAGCATGCTTGGC 714  
 Db 1865 GAACGCTAGAGAC-----GAGCTACAGAGAGATGTTCGGCTACCCATCGGCTGAT 1815  
 QY 715 TAAATTTGCAATCGTCTTGTAGATGATG-----GCGGTTGGCAATTTTCAAAAATACG 768  
 Db 1814 AAAAGCATCATCCCTCCAGTAGATGAAGCAATCCTCGGTTGTGATTCGAGTGAATTC 1755  
 QY 769 CAATTT--TTGTGCCAAATATATCCACCANTGCCATCGTCCGATCAAGATGATCAGCGGA 825  
 Db 1754 CATTTCGCTTCTGCAATAAGCTCGAATCGGTGATCGATCTAATGATCCGGTGT 1695  
 QY 826 GAGATTTAAATTTGTGCCCCCTTGAGTGGCAATTTGAGATATGCTCAAGCTGAATACT 885  
 Db 1694 AATATTAAGCAGAAATGGCACATTAAGCCCTGAAGTCTGACATGTTGTCCAGTTGAATACT 1635  
 QY 886 TGAAGCTCAAGCACCACCAATCCATATTTTGGATCGTTAATAATCAAGTGCAGGCGT 945  
 Db 1634 GCTCAACTCTATTAATATTAAGATGAAGATCGTAAGCCGACTGTCTGGCCAGACTAAA 1575  
 QY 946 GCCAATTTTACCGCCGACACCAACATCATGCTGCATGTTTGGCAATTCGCTACTAA 1005  
 Db 1574 ACCGACATTTGCTGCTCCAAACCTACATCAAAACCGGCTTACAGAGCGTGTATCAACCA 1515  
 QY 1006 TGTGTAAACAGTCTTTTGGATTTGAGCTGTGTATGCGATGATTTGTGTGTGT 1062  
 Db 1514 CATGTTGTGTGTCTTGCATTACTACCGGTATGACACCATTAATGATCGGT 1458

Search completed: May 12, 2006, 13:05:09  
 Job time : 295 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 13:01:47 ; Search time 1245 Seconds  
(without alignments)  
9166.058 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380  
Sequence: 1 taaagcgtgacacatgttaa.....acgacgcatatttttga 1380

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications NA Main: \*  
1: /cgn2\_6/ptodaca/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodaca/1/pubpna/US08\_PUBCOMB.seq: \*  
3: /cgn2\_6/ptodaca/1/pubpna/US09\_PUBCOMB.seq: \*  
4: /cgn2\_6/ptodaca/1/pubpna/US10A\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodaca/1/pubpna/US10B\_PUBCOMB.seq: \*  
6: /cgn2\_6/ptodaca/1/pubpna/US10C\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodaca/1/pubpna/US10D\_PUBCOMB.seq: \*  
8: /cgn2\_6/ptodaca/1/pubpna/US10E\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodaca/1/pubpna/US10F\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodaca/1/pubpna/US11\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length  | DB ID | Description          |
|------------|-------|-------------|---------|-------|----------------------|
| 1          | 1380  | 100.0       | 1398    | 7     | US-10-282-122A-27195 |
| 2          | 1380  | 100.0       | 96109   | 7     | US-10-672-787-35     |
| 3          | 231   | 16.7        | 1344    | 7     | US-10-282-122A-8940  |
| 4          | 160.4 | 11.6        | 1314    | 3     | US-09-815-242-7090   |
| 5          | 160.4 | 11.6        | 1314    | 7     | US-10-282-122A-22236 |
| 6          | 160.4 | 11.6        | 1314    | 9     | US-10-958-216-400    |
| 7          | 160.4 | 11.6        | 11498   | 3     | US-09-754-468-44     |
| 8          | 160.4 | 11.6        | 1830121 | 7     | US-10-329-670-1      |
| 9          | 160.4 | 11.6        | 1830121 | 8     | US-10-158-865-1      |
| 10         | 160.4 | 11.6        | 1830121 | 9     | US-10-981-687-1      |
| 11         | 154.4 | 11.2        | 14324   | 9     | US-10-795-159-685    |
| 12         | 154.4 | 11.2        | 908766  | 9     | US-10-795-159-685    |
| 13         | 152.8 | 11.1        | 1281    | 7     | US-10-282-122A-25165 |
| 14         | 152.4 | 11.0        | 1314    | 9     | US-10-958-216-402    |
| 15         | 142.2 | 10.3        | 1305    | 7     | US-10-282-122A-30713 |
| 16         | 139.6 | 10.1        | 1317    | 3     | US-09-741-669-259    |
| 17         | 139.6 | 10.1        | 1317    | 7     | US-09-815-242-5932   |
| 18         | 139.6 | 10.1        | 1317    | 7     | US-10-282-122A-20229 |
| 19         | 139.6 | 10.1        | 28277   | 3     | US-09-754-468-2      |
| 20         | 123.6 | 9.0         | 1308    | 7     | US-10-282-122A-32629 |
| 21         | 117.8 | 8.5         | 1162    | 7     | US-10-282-122A-33528 |
| 22         | 117   | 8.5         | 1347    | 3     | US-09-815-242-7912   |
| 23         | 117   | 8.5         | 1347    | 7     | US-10-282-122A-30468 |

|      |       |     |        |   |                      |                   |
|------|-------|-----|--------|---|----------------------|-------------------|
| C 24 | 113   | 8.2 | 1404   | 7 | US-10-282-122A-41274 | Sequence 41274, A |
| C 25 | 111   | 8.0 | 1317   | 7 | US-10-282-122A-39003 | Sequence 39003, A |
| C 26 | 110.8 | 8.0 | 4942   | 9 | US-10-450-763-25607  | Sequence 25607, A |
| C 27 | 110.8 | 8.0 | 4944   | 9 | US-10-450-763-25607  | Sequence 25607, A |
| C 28 | 110.8 | 8.0 | 4944   | 9 | US-10-450-763-25607  | Sequence 25607, A |
| C 29 | 110   | 8.0 | 1350   | 7 | US-10-282-122A-31716 | Sequence 31716, A |
| C 30 | 109.4 | 7.9 | 1317   | 3 | US-09-815-242-9978   | Sequence 9978, A  |
| C 31 | 109.4 | 7.9 | 1317   | 7 | US-10-282-122A-39835 | Sequence 39835, A |
| C 32 | 103.8 | 7.5 | 1314   | 7 | US-10-282-122A-23400 | Sequence 23400, A |
| C 33 | 94.2  | 6.8 | 1314   | 7 | US-10-282-122A-19656 | Sequence 19656, A |
| C 34 | 92.8  | 6.7 | 2058   | 9 | US-10-450-763-8816   | Sequence 8816, A  |
| C 35 | 92.8  | 6.7 | 4831   | 9 | US-10-450-763-30346  | Sequence 30346, A |
| C 36 | 92.6  | 6.7 | 1317   | 7 | US-10-282-122A-41740 | Sequence 41740, A |
| C 37 | 89.6  | 6.5 | 3117   | 9 | US-10-450-763-4899   | Sequence 4899, A  |
| C 38 | 89.6  | 6.5 | 3117   | 9 | US-10-450-763-25044  | Sequence 25044, A |
| C 39 | 88.6  | 6.4 | 1353   | 7 | US-10-282-122A-17455 | Sequence 17455, A |
| C 40 | 84.8  | 6.1 | 640681 | 3 | US-09-790-988-1      | Sequence 17455, A |
| C 41 | 83    | 6.0 | 1359   | 7 | US-10-282-122A-21621 | Sequence 21621, A |
| C 42 | 81    | 5.9 | 1071   | 7 | US-10-282-122A-37050 | Sequence 37050, A |
| C 43 | 73    | 5.3 | 1338   | 7 | US-10-282-122A-29879 | Sequence 29879, A |
| C 44 | 70.2  | 5.1 | 1440   | 7 | US-10-282-122A-17840 | Sequence 17840, A |
| C 45 | 67    | 4.9 | 2402   | 5 | US-10-194-163-881    | Sequence 881, A   |

#### ALIGNMENTS

RESULT 1  
US-10-282-122A-27195/c  
Sequence 27195, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zykand, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,338  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 27195  
LENGTH: 1398

TYPE: DNA  
ORGANISM: Moraxella catarrhalis  
US-10-282-122A-27195

Query Match 100.0%; Score 1380; DB 7; Length 1398;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 TAAAGGCTGAACCATGTTAAACAACCTTACCGCTGCTATTAATTTTAAATGATC 60
1383 TAAAGGCTGAACCATGTTAAACAACCTTACCGCTGCTATTAATTTTAAATGATC 1324
61 CATGCTAGCAACAAGAGTGAACAAAGCAGCCCTGACGGCTTGAATGCTGCGCA 120
1323 CATGCTAGCAACAAGAGTGAACAAAGCAGCCCTGACGGCTTGAATGCTGCGCA 1264
121 TACTGCGCATCAAGTGTACCAACCTGATATTTTAAATGCTGATGCTGCTT 180
1263 TACTGCGCATCAAGTGTACCAACCTGATATTTTAAATGCTGATGCTGCTT 1204
181 GAGTAAATCTTGTCAATTAACGGGCTGCTGACCGATTAATTAACCACTATCCAC 240
1203 GAGTAAATCTTGTCAATTAACGGGCTGCTGACCGATTAATTAACCACTATCCAC 1144
241 CTGCTAATCAATATGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 300
1143 CTGCTAATCAATATGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 1084
301 CATGCTAATCAATATGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 360
1083 CATGCTAATCAATATGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 1024
361 GCGGATGTTGATCTTGTGAATCATTAATATGCTGCTGATCAATCATCAATAT 420
1023 GCGGATGTTGATCTTGTGAATCATTAATATGCTGCTGATCAATCATCAATAT 964
421 TTCAACGCAATGCTGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 480
963 TTCAACGCAATGCTGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 904
481 AAGATTGGGAGCTGCTGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 540
903 AAGATTGGGAGCTGCTGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 844
541 TTGATTTTAAATGATGATGTTGTTAAAGTTGTTTGAATGATTAATGATGATG 600
843 TTGATTTTAAATGATGATGTTGTTAAAGTTGTTTGAATGATTAATGATGATG 784
601 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
783 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 724
661 GATTAATGCTGATGTTGCTGCTGACAGCTCACTAAATCTTGAACCTTGAACCA 720
723 GATTAATGCTGATGTTGCTGCTGACAGCTCACTAAATCTTGAACCTTGAACCA 664
721 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
663 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
781 CAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
603 CAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 544
841 TGGCCCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
543 TGGCCCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 484
901 CAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
483 CAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 424
961 GACACCAACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
```

```
423 GACACCAACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
1021 TTTGGATTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
363 TTTGGATTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
1081 TGGCTTGAAGTATCAATGAACATTAATGAACATGCTGCTGCTGCTGCTGCTGCTG 1140
303 TGGCTTGAAGTATCAATGAACATTAATGAACATGCTGCTGCTGCTGCTGCTGCTG 244
1141 GCGGCGGACATATCTGCGGCTTCTTGGTTAAATGCGGCGGCTGATGATGCTGCTG 1200
243 GCGGCGGACATATCTGCGGCTTCTTGGTTAAATGCGGCGGCTGATGATGCTGCTG 184
1201 TGAAGTTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
183 TGAAGTTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124
1261 ATCTGCAAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
123 ATCTGCAAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 64
1321 CAAAAAATTCACAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
63 CAAAAAATTCACAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4
```

## RESULT 2

```
US-10-672-787-35
; Sequence 35, Application US/10672787
; Publication No. US2004006754A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELIPIA, 02561
; CURRENT APPLICATION NUMBER: US/10/672,787
; PRIOR APPLICATION NUMBER: 2003-09-26
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 96109
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-10-672-787-35
```

Query Match 100.0%; Score 1380; DB 7; Length 96109;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 TAAAGGCTGAACCATGTTAAACAACCTTACCGCTGCTATTAATTTTAAATGATC 60
11357 TAAAGGCTGAACCATGTTAAACAACCTTACCGCTGCTATTAATTTTAAATGATC 11416
61 CATGCTAGCAACAAGAGTGAACAAAGCAGCCCTGACGGCTTGAATGCTGCGCA 120
11417 CATGCTAGCAACAAGAGTGAACAAAGCAGCCCTGACGGCTTGAATGCTGCGCA 11476
121 TACTGCGCATCAAGTGTACCAACCTGATATTTTAAATGCTGATGCTGCTGCTT 180
11477 TACTGCGCATCAAGTGTACCAACCTGATATTTTAAATGCTGATGCTGCTGCTT 11536
181 GAGTAAATCTTGTCAATTAACGGGCTGCTGACCGATTAATTAATCACTATCCACCA 240
11537 GAGTAAATCTTGTCAATTAACGGGCTGCTGACCGATTAATTAATCACTATCCACCA 11596
241 CTGCTAATCAATATGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 300
11597 CTGCTAATCAATATGCTGACAGCTCACTAAATCTTGAACCTTGAACCAAGCCCA 11656
```



```
QY 301 CAATGCCAGTGTATATGCGTAACCGTGCAGGCCATCATTTGCTGCAATGCTTGA 360
Db 11657 CAATGCCAGTGTATATGCGTAACCGTGCAGGCCATCATTTGCTGCAATGCTTGA 11716
QY 361 GCCGATGTTGTTGACCTTTTGAATCATTTAAATATGCTTGTATCATATCATATTA 420
Db 11717 GCCGATGTTGTTGACCTTTTGAATCATTTAAATATGCTTGTATCATATCATATTA 11776
QY 421 TTCAAGGAGTGGGTTAGCCCTTTGAAAGCTTAAAGGTATCAAGCATTCGATCCATGGG 480
Db 11777 TTCAAGGAGTGGGTTAGCCCTTTGAAAGCTTAAAGGTATCAAGCATTCGATCCATGGG 11836
QY 481 AAGATTGGCAGCTGTGCTTAAGCAGATGAGATTAAGCATTTGATGATTTGCTTGGC 540
Db 11837 AAGATTGGCAGCTGTGCTTAAGCAGATGAGATTAAGCATTTGATGATTTGCTTGGC 11896
QY 541 TTTGATTTTAAATGATGATTTGTTAAAGTTTGTGTTTATGATATATATGCAATACC 600
Db 11897 TTTGATTTTAAATGATGATTTGTTAAAGTTTGTGTTTATGATATATATGCAATACC 11956
QY 601 ATGACCATCATCAATTAATTAATTAATGCTGCTTGTGCTTGAAGCATTTGTTGT 660
Db 11957 ATGACCATCATCAATTAATTAATTAATGCTGCTTGTGCTTGAAGCATTTGTTGT 12016
QY 661 GATATATGTTGATTTTGGGCAACGCTTGTCAAGTGTGCTCAAGCATTTGCTTAAT 720
Db 12017 GATATATGTTGATTTTGGGCAACGCTTGTCAAGTGTGCTCAAGCATTTGCTTAAT 12076
QY 721 GGCATCTGTTGTTGATGATTTGGGTTTGGCAATTTTCAAAAATGACCAATTTTGTG 780
Db 12077 GGCATCTGTTGTTGATGATTTGGGTTTGGCAATTTTCAAAAATGACCAATTTTGTG 12136
QY 781 CAATATATCAACCATGTCATGTCGTCATCAAGTATCATGCGGAGATTTTAAATTTGT 840
Db 12137 CAATATATCAACCATGTCATGTCGTCATCAAGTATCATGCGGAGATTTTAAATTTGT 12196
QY 841 TGCCCTTTGAGCTGCCAAATTTGATATGCTCAAGCTGTAAGAACTTACAGCTCAACAC 900
Db 12197 TGCCCTTTGAGCTGCCAAATTTGATATGCTCAAGCTGTAAGAACTTACAGCTCAACAC 12256
QY 901 CACCAAAATCAATTTTGTGATGCTTAATTAATCAAGTGAAGCGCTGCCAATATATCCGCC 960
Db 12257 CACCAAAATCAATTTTGTGATGCTTAATTAATCAAGTGAAGCGCTGCCAATATATCCGCC 12316
QY 961 GACACCAACATATGCTGCTGATTTTGGCATCTGCCCTAATATGCTGTAACAGTGT 1020
Db 12317 GACACCAACATATGCTGCTGATTTTGGCATCTGCCCTAATATGCTGTAACAGTGT 12376
QY 1021 TTTGGCATTTGAGCTGTGATGAGGATGATTTGGTGTGTTGCTTGAATGATTCG 1080
Db 12377 TTTGGCATTTGAGCTGTGATGAGGATGATTTGGTGTGTTGCTTGAATGATTCG 12436
QY 1081 TGCTTTGAGTGTATCAATGAACAATTTGAACATGCTGATCAAGGAATGCTTGAAGCTTT 1140
Db 12437 TGCTTTGAGTGTATCAATGAACAATTTGAACATGCTGATCAAGGAATGCTTGAAGCTTT 12496
QY 1141 GGGGGGGAACAATCTGGGCTTCTTGGGTTAATGCGGGGCTGATGATGATTCGTCG 1200
Db 12497 GGGGGGGAACAATCTGGGCTTCTTGGGTTAATGCGGGGCTGATGATGATTCGTCG 12556
QY 1201 TGAGCTTTAAGAGTTCACTGTCCAAAGCTGCCAAATATGCAATGAACGCTGTGCGGAGTTT 1260
Db 12557 TGAGCTTTAAGAGTTCACTGTCCAAAGCTGCCAAATATGCAATGAACGCTGTGCGGAGTTT 12616
QY 1261 ATGTGCCAAGGTAGATTTGGGATTTAATCATGATGATGCTAATCTTGTGACCGTATGAC 1320
Db 12617 ATGTGCCAAGGTAGATTTGGGATTTAATCATGATGATGCTAATCTTGTGACCGTATGAC 12676
QY 1321 CAAAAAATTTGACAGCTGACAGCTGACAGCTGCCAAACGACGACGACGATTTTGTGA 1380
Db 12677 CAAAAAATTTGACAGCTGACAGCTGACAGCTGCCAAACGACGACGACGATTTTGTGA 12736
```

```
RESULT 3
US-10-282-122A-8940/C
; Sequence 8940, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELPTRA 03A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8940
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-8940

Query Match 16.7%; Score 231; DB 7; Length 1344;
Best Local Similarity 52.4%; Pred. No. 2,7e-53;
Matches 721; Conservative 0; Mismatches 590; Indels 66; Gaps 7;

QY 2 AAGCGTGAACCATGTTAACAACACTATACCTGCTGCTATATATTTTAAATGATCC 61
Db 1340 AAGCAATTGACGACGACCAACAACTGCTGACAGGTCATTTATACCTTTTAAACATATCA 1281
QY 62 ATGCTGACCAAGCAGGTGACAAAGACAGCCGCTGAGCGGTGCTTGAAGCTGCGCAAT 121
Db 1280 AACTTGCACATGCTGTGAT-----AGCAATACCAATCTTGAAGCTTGTTCAC 1229
QY 122 ACTGCCGATCAAGTGTACCAACTGATATATTTTGAATGCCAATGTGCATCACCTTTG 181
Db 1228 GTTGAACACAGCTGACAGCTTCTTTAAGGCTTGTGCTGATGTAATTTAGTTCGCTT 1170
QY 182 AGTAATCTTGTCTCAATTTATGGGCGATCTTGACCATTAATATACATATCCACCCAC 241
Db 1169 TGAATGCTTGTCTGATGACCGGCGCATCTTACCAATATACCAACTTTGGCATAT 1110
QY 242 TGCATAACATATGCTGACAGCTCACTAATCTTGAACCTTTACCAAGCCCAACATC 301
```

Db 1109 TTTTCATNAGAGACGTAAGAGAAAACTTGCCCTTTGCTGCTCCACTTAATAA 1050  
QY 302 AATGCA-----GTATGATATGCCATAAACCGTGCAGACCCATCAATTTGTCGAA 355  
Db 1049 AGTGCACACCTTACTTTTTCATCTTCAATGGCAGACCTTAAGCATCAATTTGTCGAA 990  
QY 356 GTTGAAGCCGATGTTTGAATCTTTGAATCAATTAATATGCTTGTCTATCAATATCA 415  
Db 989 GTAGCAGCAACATTTGATCTTTTGAAGTCAATTAATACGACATCACACAGTTTAA 930  
QY 416 ATATATTCACGAGATGCGGTAGCCCTTGAAGTCTTAAGGTATCAACAGGACATCC 475  
Db 929 ACCTACTCAGAGAGTCTTAATCTTAAATCGTTTAAAGTTTCAAGCATTTGCTCC 870  
QY 476 ATGGAGAGATTGACAGTGTGCTTAAGGCAAGTCAAGATTAAGCATTTAGTAGTATGC 535  
Db 869 ATAGTAACCAATTTGCTCTCCCAATGCTAAACAGCTTAACATTTGCTAGTGTGC 810  
QY 536 TTGCTTTGATTTTAAATGATGATTTGTAAGTTGTTTGTGATGATTAATGCC 595  
Db 809 ATACCTGATATATTAATCCGAGCTTTTAAATACCGCTGTAACACAGACAGCA 750  
QY 596 ATACCATCACCATCAATAATAATAAAATCGCGTTTGTGCTAAGTATGATTT 655  
Db 749 AGCGTACCGTGGCATCTCTTAACACGCAATTTGATTT-----TAAATCTGTGCGT 698  
QY 656 GTTGTATATGTTGATTTGGGCAACCGTTGTCAAGTGTGCAAGCTTGTGCT 715  
Db 697 TTAAACCAAGCTTTGATTTGATGATGATGATGATGATGATGATGATGATGAT 647  
QY 716 AAATTGGCATGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 775  
Db 646 ---AAGCATGCTGCTGTTAAATACAACTTTTGAAGCGCTTGAAGAAATACATGTT 591  
QY 776 TGTGCAAAATATCAACCATGATGCTGCGATGCAAGATGATGATGATGATGAT 835  
Db 590 GCTTGTGATACCCAGCATATTTTCAATGAGGCTTAATGATGATGATGATGATGAT 531  
QY 836 ATTTGTGCTGCTTGAAGTGTGCAATTTGATGATGATGATGATGATGATGATGAT 895  
Db 530 ACCAGTCTACCTCAAGGTTTAAAGTGTGATGATGATGATGATGATGATGATGAT 471  
QY 896 AGACACCAAAATTCATATTTTGAATGATGATGATGATGATGATGATGATGATGAT 955  
Db 470 AGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 414  
QY 956 CCGCCGACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015  
Db 413 CGGCTACTGCACTTTTCTTACCTGATCTTACGATGATGATGATGATGATGATGAT 354  
QY 1016 GTGCTTTTGGCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1075  
Db 353 GTATCTTTTGGCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 304  
QY 1076 TCTGCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1135  
Db 303 -----ACGACGCAATTAATGATGATGATGATGATGATGATGATGATGAT 261  
QY 1136 GCTTTGGCGGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1195  
Db 260 GCAATGCTGCTGCAATTTCCGCTTAATGATGATGATGATGATGATGATGATGAT 201  
QY 1196 TCGGCTGAGCTTAAGATTTCACTGTGCAAGCTGCAAAATGACATGAAAGCTGTGGC 1255  
Db 200 TCTGCTGATTAATTAATTTTATTTGATCAAGCTGCAAAATGCTGTTTAAAGCGGCA 141  
QY 1256 AGTTATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315  
Db 140 ATCTGATGCTGCTCGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 81  
QY 1316 TTGACCAAAATTTTCAAGCTGACATGATGATGATGATGATGATGATGATGATGAT 1372  
Db 80 TCAATGAGAAATTTTACAGCAAAACCTGATATTTCCAGCTGCTGTAACCTTTT 24

RESULT 4  
US-09-815-242-7090/c  
Sequence 7090, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl U.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7090  
LENGTH: 1314  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) ... (1314)  
US-09-815-242-7090  
Query Match 11.6%; Score 160.4; DB 3; Length 1314;  
Best local similarity 49.6%; Pred. No. 1.4e-33;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
QY 1 TAAAGGTGAACCATGTTTAACAACCTATACCTGCTCATTAATTTTAAATGATC 60  
Db 1308 TAAACATTAAGCTTAATGCTTAATTTCTGCGGCTTTTCAAAAGAAAGCATGATC 1249  
QY 61 CATGCTAGCAACAGAGGTGACAAAAGCAGCCCTTGAAGGCTTGAATGCTGCCAA 120  
Db 1248 GAGACTTGCACAGAGGCGGACAAATTAATCA-----TATCTCCCTTTGCA 1201  
QY 121 TACTGCCGATCAAGTGTACCAACCTGATATATTTAATGTCGAATGATGATCCTT 180  
Db 1200 TGTGGGCTTAATAATTTCTATGCTTTGTTCCATTTGATGAACA-----ATACTTTG 1147  
QY 181 GAGTAATCTTCTCAATTTAGGGGCACTTGAACGATTAATAACAATATCAACCA 240  
Db 1146 CGATGAATTTTGGACAGAGGCAACATCTGACCAAAACAATTAATATGATG 1087  
QY 241 CTGCATTAATATGCTGACAGCTCACTTAATCTTGAACCTTTAACAAGCCCAAAAT 300  
Db 1086 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1027  
QY 301 CAATGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Db 1026 CAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 976

```

QY 361 GCCGATGTTTACCTTTGATGATTAATAATGCTTGTATCATATCATCAATATA 420
Db 975 CCCCAATTTGTTGTTAGAGTCAATTAATCAACGAATGCAATAGCTTATGACATA 916
QY 421 TTCACAGGATGCGGTACCTTTGAAAGCTTAAGGTTACAGATGSCATCAGGG 480
Db 915 TTGAAACGATGATCACTTTGAAATGACGAAGTCCGATGCAATGTAATCTAAAT 856
QY 481 AAGATGCGAGCTGCTTAGGCAAGGAGATGAGCATGATGATGATGCTTGGCC 540
Db 855 AATACCTATGCTTGTGCTCAATGCTGTGCTCCAAATGTTCAATATATATGCGACC 796
QY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 795 AACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 736
QY 601 ATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 735 GCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
QY 660 TGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
Db 675 CGCTGATTTTCTGTTTCCCAAAATGACGC----- 644
QY 720 TGGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 643 ---TATCTTATGTTCAACACACTTATGATGATGATGATGATGATGATGATGAT 587
QY 780 CCAATATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 586 GGCATATATCTTCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 527
QY 840 TTGCCCCCTGAGCGCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 899
Db 526 TCGTGTGCGAGCTTTTAAGCTATATGATGATGATGATGATGATGATGATGATGAT 467
QY 900 CCAACCAATCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
Db 466 CATTAAGTTCAATCTTCAATTC---AACAATGACAAAGCGGAAATCCATATATTTCCG 410
QY 960 CGACACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
Db 409 CCAATACCACTTTCACACGAGAGCTTCCCAATTCATTAATGATGATGATGATGAT 350
QY 1020 TTTTGGATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
Db 349 TTTTACCATTTGAACTGTAATCCCAATGCTTGTGCTGCGCGCA----- 298
QY 1080 GTGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139
Db 297 -----GATTAATTCATATGCGGATTAATCTTCCATCCGCTTAA 257
QY 1140 TGGCGGCGCAATCTCGCGGTTTGGTTAATGCGGCGCTGATGATGATGATGATGAT 1199
Db 256 GTGCGGTTTGAATTTCTGTTGTTTACCGCAAGCCCTGGGCTTAATCAATCATATGCG 197
QY 1200 CTGAGCTTAGAGTCACTGTCAGAGTCCAAATGACAAATGAAACCCCTGTGGCGATT 1259
Db 196 TTTTCAAGTATCAATTTCTGATTTTAACCTACAGTATGAAAGGATATTTTGAAGAAATT 137
QY 1260 TATCTGCGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319
Db 136 TATCAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77
QY 1320 CCAAAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
Db 76 AGAGATATATCACAAGAAGAGCTGTTTGGCAAGCCGATGATGATGATGATGATGAT 19

```

RESULT 5  
 US-10-282-122A-22236/c  
 ; Sequence 22236, Application US/10282122A  
 ; Publication No. US20040029129A1

```

: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari
: APPLICANT: Zykkind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: APPLICANT: Xu, H.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: EPIITA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: PRIOR FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 78614
: SOFTWARE: PatentIn version 3.1
: LENGTH: 1314
: SEQ ID NO 22236
: TYPE: DNA
: ORGANISM: Haemophilus influenzae
: US-10-282-122A-22236

```

Query Match 11.6%; Score 160.4; DB 7; Length 1314;  
 Best Local Similarity 49.6%; Pred. No. 1.4e-33;  
 Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```

QY 1 TAAAGGTGAACCATGTTAACAACCTTATCATCTGCTCATTAATAATTTTAAATGATC 60
Db 1308 TAAACATTTGAGCTTAATGCTTAATTTCTTCCGCGCTTTTCAAAAGCAATGATC 1249
QY 61 CATGCTAGCAAGAGGTGACAAAGCAAGCCCTTGAAGGCTTGAAGCTGCGCA 120
Db 1248 GAGACTTGGCAAGAGGCGCAATATACCA-----TATCTCGCTTGGCA 1201
QY 121 TACTGCCGATCAAGTGTACCACTGATATATTTTGAATGCAAAATGTCATCAGCTTT 180
Db 1200 TGTGGGCGTAAATAATTTATGCTTGTTCATTTGATGAAACA-----ATAACTTGG 1147
QY 181 GAGTAATCTTGTCAATTAATGCGGCAATCTTGAACGATTAATATACACTATCCACCA 240
Db 1146 CGATGAAATTTTGGCAAGAGCGCACATCTGCACAAAACAATATATATGTGTG 1087
QY 241 CTGCAATACATATGCGAGAGCTCACTAAATCTTGAACCTTAAACCAAGCCCAAAAT 300
Db 1086 TTGATTTAATTAATTCAGCTTAATTTGAAATAATGAGCCCTTTTGGCTTCCGCTAGCA 1027
QY 301 CATGCGAGTATGATATGCAATTAACCGTGCAGGCCCATCAATTTGCTGCATGATGTTGA 360
Db 1026 CAATGCAATTTAACCTCAATATA-----AGCCAGGCAATGACGCAACTTACT 976

```

QY 361 GCCGATGTTGTAACCTTTGAAATCATTAATAATGCGCTGATCATATATCATCAATATA 420  
DB 975 CCCCAATTTGTTGCTTTAGATCATTAATCAACGAATGCATTTAGCTGATGACATA 916  
QY 421 TTACAGCGATGGGATGACCTTTGAAAAGCTTAAGGATTCAGACATGCAATTCATGGG 480  
DB 915 TTAAAAACGATGATCTTAACCTTTGAAATGACGAAGGCGGTGCAATTTCAATTAATT 856  
QY 481 AAGATTGGCAGTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGAGTATGCTTATGCTGCC 540  
DB 855 AATACCTATGCTTGTGCGCAATGCTGTGCTGCCAAATGTTCAATTAATTAATGCGACCC 796  
QY 541 TTTGATTTTAAATGATGATGATTAAGTTGTTTGTGTTGATTAATATGCAATACC 600  
DB 795 AACCAATGATGCTTCTTCAACAGTAAATACACTTCACTTTTAAACATTAATATGCTT 736  
QY 601 AATACATATCAAAATTAATTAATAATATGCGCTTTGATGCTTAAGTATGCTTATGCTG 659  
DB 735 GCCATTTTCAATTTTAAAGCAATTAATCCGCACTAATTTTCCGCAAAAGAAACGGTATGTT 676  
QY 660 TGAATTAATGATGATTTGGGCAACGCTTGTCAATGATGATCAAGCATGCTTGGCTAAT 719  
DB 675 CGCTTGAATTTTCTTTTCCCAAAAGTCAAGC----- 644  
QY 720 TGGCATGCTTGTGATGATGATGCGGCTTTGGCAATTTTCAAAATACGCAATTTTGTG 779  
DB 643 ---TATCTTCAATGTTCAACACACTTACTTACATTAATGATTAATGCGTAATTTTGCTT 587  
QY 780 CCAATTAATCCACATGCGCATGCGCATCAAGATGATGAGGAGAGATTTAAATTTG 839  
DB 586 GCGGATTAATCTTTAAATCCATTAATGATGATGATGATGATGATGATGATGATGATGAT 527  
QY 840 TTGCGCTTGAAGTGCCTTGAATTTGAGATATGCTCAAGCTAAATTTGACGCTCAAGCA 899  
DB 526 TCGCTGCGGAGCTTTTAAGCTTAAGTGTCTCAAGCTTAAGCTTAAGCTTAAGCTTAAGCT 467  
QY 900 CCAACCAATCAATATTTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 959  
DB 466 CATTAAGTTCAACATCTTCACTTC---AACATGACAAAGGGGAATCCCAATATTTCCGG 410  
QY 960 CGAACCAACATCAATGCGCTGATGATTTTCCATCTGCTTCAATATGCTTAAGCTGACGTC 1019  
DB 409 CCAATCAACCTTTGACACACGAGCTTTGCGCAATTTCAATTAATTAATTAATTAATTAAT 350  
QY 1020 TTTTGGCATTTGAGCTGTGATGAGGATGATGATGATGATGATGATGATGATGATGAT 1079  
DB 349 TTTTACATTTTGAACCTGTAATCCCAATTTGCTTTGCTGCTGCGGCA----- 298  
QY 1080 GTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139  
DB 297 -----GAATTAATTAATTAATGCGCTGATGATGATGATGATGATGATGATGATGAT 257  
QY 1140 TGGCGGCGCAATCAATCTGCGCTTCTGCTGATTAATGCGGCGCTGATGATGATGATGAT 1199  
DB 256 GTGCGGCTTGAATTTCTGCTGTTTAAACGCAAGCCCTGCGCTTAATTAATCAATATATCC 197  
QY 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
DB 196 TTTTCAAGTAACTTCTGATTTTAACTTAAGTAACTTAAGTAACTTAAGTAACTTAAGTAA 137  
QY 1260 TATCTGCGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319  
DB 136 TATCATATACAGATGATGATTTTTCGAGTATCAATCAACGAATATTAAGCTGTTGGGATA 77  
QY 1320 CCAAAAAATTCACAGCTGACATGCGCAAGCTGCGCAACCGACGAGGCAATTTT 1377  
DB 76 AAGATTAATTCACACAAAGAAACCTGTTTTCGCAAGCCCGATGATGATTAATTTT 19

RESULT 6  
US-10-958-400/C  
; Sequence 400, Application US/10958216

Publication No. US20050181388A1  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, ALBED  
APPLICANT: DHARMSI, AKIL  
APPLICANT: VEDADI, MASOUD  
APPLICANT: ALAM, MOHAMMAD ZAHOOR  
APPLICANT: ARROWSMITH, CHERYL  
APPLICANT: AMREY, DONALD E.  
APPLICANT: BEATTIE, BRYAN  
APPLICANT: BUZADZILA, KRISTINA  
APPLICANT: CANADIER, VERONICA  
APPLICANT: DOMAGALA, MEGAN  
APPLICANT: HOUSTON, SIMON  
APPLICANT: KANAGARAJAH, DHUSHY  
APPLICANT: LI, QIN  
APPLICANT: MANSOURY, KAMRAN  
APPLICANT: MCDONALD, MERRY-LYNN  
APPLICANT: NETHERY, KATHLEEN  
APPLICANT: NG, IVY  
APPLICANT: OUYANG, HUI  
APPLICANT: PINDER, BENJAMIN  
APPLICANT: RICHARDS, DAMN  
APPLICANT: TAI, MATTHEW  
APPLICANT: THALAKADA, ROSANNE  
APPLICANT: VALLEE, FRANCOIS  
APPLICANT: VIRAG, CRISTINA  
TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM BACTERIA  
FILE REFERENCE: IPT-205.01  
CURRENT APPLICATION NUMBER: US/10/958,216  
CURRENT FILING DATE: 2004-10-04  
PRIOR APPLICATION NUMBER: PCT/CA03/00462  
PRIOR FILING DATE: 2003-04-02  
PRIOR APPLICATION NUMBER: 60/369,511  
PRIOR FILING DATE: 2002-04-02  
PRIOR APPLICATION NUMBER: 60/385,089  
PRIOR FILING DATE: 2002-05-31  
PRIOR APPLICATION NUMBER: 60/385,751  
PRIOR FILING DATE: 2002-06-04  
PRIOR APPLICATION NUMBER: 60/386,553  
PRIOR FILING DATE: 2002-06-05  
PRIOR APPLICATION NUMBER: 60/386,577  
PRIOR FILING DATE: 2002-06-05  
PRIOR APPLICATION NUMBER: 60/386,367  
PRIOR FILING DATE: 2002-06-05  
PRIOR APPLICATION NUMBER: 60/386,566  
PRIOR FILING DATE: 2002-06-05  
PRIOR APPLICATION NUMBER: 60/386,390  
PRIOR FILING DATE: 2002-06-06  
PRIOR APPLICATION NUMBER: 60/386,601  
PRIOR FILING DATE: 2002-06-06  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1132  
SOFTWARE: PatentIn Ver. 3.3  
SEQ ID NO 400  
LENGTH: 1314  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
US-10-958-216-400

Query Match 11.6%; Score 160.4; DB 9; Length 1314;  
Best Local Similarity 49.6%; Pred. No. 1,4e-33;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
QY 1 TAAAGCGTAACCAATGTTAAACCAATTAATCACTGCTCATTAATTAATTTTAAATGATC 60  
DB 1308 TAAACATTAAGCTTAATGCTTAATTTCTTGGCGGCTTTCAAAAGAAACCAACTGATC 1249  
QY 61 CATGCTAGACACAAGAGGTGACAAAGCAAGCCCTTAAGCGGTCTTTGACTGCTGCCAA 120  
DB 1248 GAGACTTGACACAAGAGGAGGCAATTAATCA-----TATCTCCGCTTTGCA 1201  
QY 121 TACTGCCGATCAAGTATCAACCTGATTAATTTTAATGATGCCAAATGTCATCACTTT 180

Db 1200 TGTGGCGTAAAAATTCATCGCTTGTTCATGTATCGAACAA-----ATACTTGG 1147  
Qy 181 GAGTAAATCTTGTCAATTAATCGGGGATCTTGACCGGATTAATATACAGATATCCACCA 240  
Db 1146 CGATGAAATTTTGGACAGCAGCGACATCTCGACCAAAACATTAACAAATATGTGTGG 1087  
Qy 241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTTCAGCTTTTACCAAGCCCAACCAAT 300  
Db 1086 TTGATTAATTAATTCAGCTAATCTGAATAATACGCCCTTTTCGATCGCTGACCA 1027  
Qy 301 CAATGCCAGATGATATGCTTAACCGTCCAGCCCAATCAATCTGCAATGCTTGA 360  
Db 1026 CAATGCAATTTACCTCAATATA-----AAGCCCAAGCCCAATCAGCACTGTACT 976  
Qy 361 GCCGATGTTGTACCTTTGATTAATTAATATGCTTGTATCATATCATATATATA 420  
Db 975 CCCCACTTTGTTGCTTTAAGTATTAATTCAGCAAGATGCAATGCTTGTATGACATA 916  
Qy 421 TTGACAGCGATGGGATGAGCTTTGAAAGTCTTAAGGGTATCAAGCATGCGCATGAG 480  
Db 915 TTGAAACGATGATCTAACCTTTGAATGACGAAGCGGTAGCATTTGATTAATTAAT 856  
Qy 481 AAGATTGGACGCTGTGCTTAAGCGAGTGAATAGGCAATTGATGATGTTATGCTTGC 540  
Db 855 AATACCTATGCTGTGCGCAATGCTGTGCGCAAAATGTTATTAATTAATGCGAC 796  
Qy 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 795 AACCAATGATGCTTCTTCAAGGATTAATCACTTCACTTTTACATTAATTAATGCTT 726  
Qy 601 ATGACCATCATCAATTAATTAATTAATCTGC-CGTTGGTGGCTTAAGTATGCTATTTGG 659  
Db 735 GCCATTTTGAATTTTATGCAATTAATCGCATATTTTCCGCAAAAGAAAGGATGCTT 676  
Qy 660 TGAATATGCTGATTTGGGCAACGCTGTCAATGCTGATGATGATGATGATGATGAT 719  
Db 675 CGCTGATTTTCTTTTCCCAAAAGTCAAGC----- 644  
Qy 720 TGGCATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779  
Db 643 ---TATCTTATGCTTCAACACATCTTATGATTAATTAATTAATGCTTATTTGGCT 587  
Qy 780 CCAATTAATCACAATGCTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 839  
Db 586 GGGATTAATCTTAAATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 527  
Qy 840 TTGCCCCCTGAGCTGCAAAATTTGATATGCTCAAGCTGCAAAATTTGATGATGATGAT 899  
Db 526 TCGCTGCGGACGCTTTTAAAGCTAATGATGCTCAAGCTGAAATGATGATGATGATGAT 467  
Qy 900 CCAACCAATTCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959  
Db 466 CATTAAGTTCACATCTTCAATC---AACAATGACAAAGGGGAATCCCAATTTTCCG 410  
Qy 960 CGACACCAACATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
Db 409 CCAATCAACTTTCACACGACGCTTTCGCACTTTCATTAATTAATTAATTAATTAATTA 350  
Qy 1020 TTTTGGCATTTTGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
Db 349 TTTTACATTTGAACTGTATATCCCAATTTGCTTGTGCTGCGGCGCA----- 298  
Qy 1080 GTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1139  
Db 297 -----GAATTAATTCATTAATGCGGATTAATCTTCACTCCGCTTTTA 257  
Qy 1140 TGGCGGCGACATATCTGCGCTTCTGGGTTAATGCGGGCTGATGATGATGATGATGATGAT 1199  
Db 256 GTGCGGTTGAATTTTGTGCTTTTAAACGCAAGCCCTGGCTTAATTAATTAATTAATTA 197  
Qy 1200 CTGACCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
Db 196 TTTCAAGTAACTTCTGATTTAACTACAGTATGAAAGGATATTTTGAAGAGTT 137

Qy 1260 TATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319  
Db 136 TATCAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77  
Qy 1320 CCAAAAAATTCACAGCTGACATGCTCAGAGCTGCCAAGCGACGACGATATTTT 1377  
Db 76 AGATATATTCACACAAAGAAAGCTGTTTGTGCAAGCCGATGATGATGATGATGAT 19  
  
RESULT 7  
US-09-754-468-44/c  
; Sequence 44, Application US/09754468  
; Publication No. US20050192237A1  
; GENERAL INFORMATION:  
; APPLICANT: Iversen, Patrick L.  
; TITLE OF INVENTION: Antisense Antibacterial Cell Division  
; FILE REFERENCE: 0450-0033.30  
; CURRENT APPLICATION NUMBER: US/09/754,468  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: US 60/174,484  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 11498  
; TYPE: DNA  
; ORGANISM: Hemophilis influenza  
US-09-754-468-44

Query Match 11.6%; Score 160.4; DB 3; Length 11498;  
Best Local Similarity 49.6%; Pred. No. 5e-33;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy 1 TAAAGGTGAACCATGTTAAACAACTTATCACCTGCTCATTAATTAATTTTAATGATC 60  
Db 9122 TAAACATTAAGCTTAATGCTTAATTTCTGCGGCTTTCAAAAGAAAGCAATGATC 9063  
Qy 61 CATGCTAGCAACAGAGTGAACAAAGCAAGCCCTTGAAGCTGCTGCA 120  
Db 9062 GAGACTTGAACAAGCGGACGACATTAATACCA-----TATCTCCGCTTGGCA 9015  
Qy 121 TACTGCCGATCAAGTGAACCACTGATATATTTAGTGCATATGTCATGCTT 180  
Db 9014 TGTGGGCGTAAATAATTAATGCTGCTTGTCAATGTATGAACAA-----ATACTTTG 8961  
Qy 181 GAGTAAATCTGCTCAATTAATGCGGACATCTTGAACGATTAATACACTATCCACCA 240  
Db 8960 CGATGAAATTTTGGACAGCGGACACATCTGACCAAAACATTAACAAATATGTGTGG 8901  
Qy 241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTTACCAAGCCCAACCAAT 300  
Db 8900 TTGATTAATTAATTAATGCTTAATTTGAAATTAATGAGCCCTTTTCCGCTCCGCTAGCA 8841  
Qy 301 CATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
Db 8840 CAATGCAATTTTAACTCAATATA-----AAGCCCAAGCAATGACGACCACTGTACT 8790  
Qy 361 GCCGATGTTGTACCTTTTGAATCATTAATATGCTGCTGATCAATCAATATATA 420  
Db 8789 CCCCACTTTTGTGCTTTAAGTATTAATTCAGCAAGATGCAATTAATGCTTATGCACTAA 8730  
Qy 421 TTGACAGCATGCGGATGAGCTTTTGAAGTCTTAAGGATATCAAGCATGCAATGCAATG 480  
Db 8729 TTGAAACGATGATCTTAACCTTTTGAATGACGAAGTGGGATAGCAATTAATTAAT 8670  
Qy 481 AAGATTGGAGCTGTGCTTAAGGCAAGTGCATTAAGCATTAAGATGATGATGATGATGAT 540  
Db 8669 AATACCTATGCTGTGCGCAATGCTGTGCTGCAAAATGTTCAATTAATTAATGCGAC 8610  
Qy 541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

Db 8609 AACCAATGAGCTTCTTCA CAAAGTAAATCACTTATCTTTTACATTAATTAATGCTT 8550

Qy 601 ATGACCATCATGAAATTAATTAATCTGC -CGTTGGGCTAGTAAGCTATTTGTTG 659

Db 8549 GCCATTTTCAAGTTTATGCAATTAATCCGACATATTTTCCGAAAGAAACGGTATGTTT 8490

Qy 660 TGATTAATGGTGTGATTTGGGGCAACGCTGTGAGTGTGGTCAAGCATTTGCTGCTAAT 719

Db 8489 CGCTTGATTTTGGTTTTTCCCAAAAGTCAGCC----- 8458

Qy 720 TGGCATCGTCTTGATGATGATGCGGCTTTGGCAAATTTCAAAATAGCAATTTTGTG 779

Db 8457 ---TATCTTCAATGTTTCAACACACTACTTTAGCATTTGATTAATTTGGTTT 8401

Qy 780 CCAATTAATTCACACATGCGATCGTCCGATCAAGATGATCAGCGAGATTTAAATCTG 839

Db 8400 GCGCATATCTTTTAATTCATATAGCGATCATATGATCTTCACTACGTTCAAGACAG 8341

Qy 840 TTGCCCCCTTGAGCTCCAAATTTGAGATATGCTCAAGCTGAAACCTTGAACAGCTCAGCA 899

Db 8340 TCGCTGCGCGAGCTTTTAAGCTATTAAGTTGTCTCAAGCTGAAATCAGAAAGCTCTAGTA 8281

Qy 900 CCACCAATTCATATTTTGAATCGTTAATTAATTCAGTGCAGCGCTGCCAATATTAACGC 959

Db 8280 CATTAAGTTTCAACATCTTCATTC---ACAAAGACAAAGCGGGAATCCCAATATTTCCGC 8224

Qy 960 CGACACCAACATCATGCTGCGATGTTTGGCATCTGCTCTACTATATGCTGAACAGTGC 1019

Db 8223 CCATACCAACTTTGACACACAGAGCTTTGCCATTTCAATTAATTAAGTATACGTAC 8164

Qy 1020 TTTTGGCATTTGAGCTGTGATGCGGATGATGTGTGTGTTGCTTGGATGATGATCTC 1079

Db 8163 TTTTACCATTTGGAACCTGTAAATCCCAATTTGGCTTTGTGCTCTCCGGCGCA----- 8112

Qy 1080 GTGCTTGAATGATCATTAATGAACATTTGAACATGCTGATCAAGAAATGCTTGAGCTT 1139

Db 8111 -----GAATATTCATATTCGCGATTAATCTTCCACTCCGCTTTAA 8071

Qy 1140 TGGCGGCGACAAATCTCGCGCTTCTTGGTTAATCCCGGGGCTGATGATCTGATCGG 1199

Db 8070 GTGCGGTTTGAATTTCTGATGTTTATCCGCAAGCCTCGGGCTAATTAACATCATATCGC 8011

Qy 1200 CTGAGCTTAAGATTCAGTGCATGTCAGCTGCCAAATGACAAATGACGCTGTGGGCACTT 1259

Db 8010 TTTCAAGTAACCATTTCTGATTTTAACATCAAGTATGAAGGATATTTTGAAGGAAGTT 7951

Qy 1260 TATCTGCAAGTAGATTTGGGATTAACATCGATGATGATCTAATTTTGAACCGTATGA 1319

Db 7950 TATCAATACAGTAGATTTTTCGATATCAATCAACGAATATTACCTGTTGGGATA 7891

Qy 1320 CCAAAAAATTCACAGCTGACAGCTCCAGAGCTGCCAAACCGACGAGCGCATATTTT 1377

Db 7890 AGAGATTAATCCACACMAAGAACCTGTTTTCGCAAGCCGATGATTAATATTTT 7833

RESULT 8

US-10-329-670-1/c

; Sequence 1, Application US/10329670

; Publication No. US20040018503A1

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag

; FILE REFERENCE: P8186P1

; CURRENT APPLICATION NUMBER: US/10/329,670

; CURRENT FILING DATE: 2002-12-24

; PRIOR APPLICATION NUMBER: US 09/643,990

; PRIOR FILING DATE: 2000-08-23

; PRIOR APPLICATION NUMBER: US 08/487,429

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/426,787

; PRIOR FILING DATE: 1995-04-21

; NUMBER OF SEQ ID NOS: 1

```

1 SOFTWARE: PatentIn version 3.1
2 SEQ ID NO 1
3 LENGTH: 1830121
4 TYPE: DNA
5 ORGANISM: Haemophilus influenzae
6 FEATURE:
7 NAME/KEY: misc feature
8 LOCATION: (4747)..(4747)
9 OTHER INFORMATION: n equals a, t, g or c
10 FEATURE:
11 NAME/KEY: misc feature
12 LOCATION: (9921)..(9921)
13 OTHER INFORMATION: n equals a, t, g or c
14 FEATURE:
15 NAME/KEY: misc feature
16 LOCATION: (10150)..(10150)
17 OTHER INFORMATION: n equals a, t, g or c
18 FEATURE:
19 NAME/KEY: misc feature
20 LOCATION: (29258)..(29258)
21 OTHER INFORMATION: n equals a, t, g or c
22 FEATURE:
23 NAME/KEY: misc feature
24 LOCATION: (36543)..(36543)
25 OTHER INFORMATION: n equals a, t, g or c
26 FEATURE:
27 NAME/KEY: misc feature
28 LOCATION: (36551)..(36551)
29 OTHER INFORMATION: n equals a, t, g or c
30 FEATURE:
31 NAME/KEY: misc feature
32 LOCATION: (36636)..(36636)
33 OTHER INFORMATION: n equals a, t, g or c
34 FEATURE:
35 NAME/KEY: misc feature
36 LOCATION: (40808)..(40810)
37 OTHER INFORMATION: n equals a, t, g or c
38 FEATURE:
39 NAME/KEY: misc feature
40 LOCATION: (44416)..(44416)
41 OTHER INFORMATION: n equals a, t, g or c
42 FEATURE:
43 NAME/KEY: misc feature
44 LOCATION: (44905)..(44905)
45 OTHER INFORMATION: n equals a, t, g or c
46 FEATURE:
47 NAME/KEY: misc feature
48 LOCATION: (44975)..(44975)
49 OTHER INFORMATION: n equals a, t, g or c
50 FEATURE:
51 NAME/KEY: misc feature
52 LOCATION: (45593)..(45593)
53 OTHER INFORMATION: n equals a, t, g or c
54 FEATURE:
55 NAME/KEY: misc feature
56 LOCATION: (45732)..(45732)
57 OTHER INFORMATION: n equals a, t, g or c
58 FEATURE:
59 NAME/KEY: misc feature
60 LOCATION: (47036)..(47036)
61 OTHER INFORMATION: n equals a, t, g or c
62 FEATURE:
63 NAME/KEY: misc feature
64 LOCATION: (51334)..(51334)
65 OTHER INFORMATION: n equals a, t, g or c
66 FEATURE:
67 NAME/KEY: misc feature
68 LOCATION: (51602)..(51602)
69 OTHER INFORMATION: n equals a, t, g or c
70 FEATURE:
71 NAME/KEY: misc feature
72 LOCATION: (51786)..(51786)
73 OTHER INFORMATION: n equals a, t, g or c

```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (139510)..(139510)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

Query Match 11.6%; Score 160.4; DB 7; Length 1830121;  
Best Local Similarity 49.6%; Pred. No. 9.3e-32;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
QY 1 TAAAGCGTGACACATTTAAACAACCTTACCTGCTCATTAATTTTAAATGATC 60
DB 1204439 TAAACATTAGCTAAATGCGTAAATCTTGCGCGCTTTCAAAAGAACTGATC 1204380
QY 61 CATGCTAGCAACAAGCTGACAAAGCAACGCCCTAGCGGTGACTGCTGCCAA 120
DB 1204379 GAGACTTGCAACAGCGGAGCAATATACCA-----TATCTCCGTTGGCA 1204332
QY 121 TACTGCCGATCAAGTGAACCACTGATATATTAGATGCCAAATGTCATCACCTT 180
DB 1204331 TGTTCGGCGTAAATAATCTATCGCTTGTCATTGATGAACAA-----ATACTTGG 1204278
QY 181 GAGTAATCTGCTCAATTAATGCGGCACTTGAACGATTAATAACACTATCCACCA 240
DB 1204277 CGATGAATTTTTCAGAGCGCACATCTGACAAACATTAACAATATATGTGG 1204218
QY 241 CTGCATACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT 300
DB 1204217 TTGATTAATTAATGCTAATTTGAAABAATCAGCCCTTTCCGTCCTGAGCA 1204158
QY 301 CAATGCCAGTATGATATCCATTAACCGTGCCAGGCCATCAATTTGCTGCAATGGTTA 360
DB 1204157 CAATGCCAATTTACCTCAATATA-----AAGCCAGCCAAATGACGAACCTGACT 1204107
QY 361 GCCGATGTTTGAACCTTTGAATCATTAATATATGCTTGATCATATATCATATATA 420
DB 1204106 CCCACATTTGTTGCTTATGAGTCATTAATCCAAAGAAATGCCATTAAGCTGACCTTA 1204047
QY 421 TTCACAGCATGCGGTAGCCCTTTGAAGTCTTAAAGGATATCAAGCATGCAATCCATGG 480
```



|    |         |   |         |
|----|---------|---|---------|
| Db | 1204046 | TTGAAAACGATATCTAACCCCTTTGAATGAGAAAGTCGGATACGAATTTGAATCTAAAT     | 1203987 |
| Qy | 481     | AAGATTGGCAGCTGTGGCTTAAGGCAAGTCAGATTAAGCATTAGAGTATATGCTTGC       | 540     |
| Db | 1203986 | AATACCTATAGCTTGTGGCAATGCTGTGTGCTGCCAAAATGTTCATATATATTTATGCGACC  | 1203927 |
| Qy | 541     | TTTGATTTTAAATGATGGAATTGTAAAGTTGTGTTTTTGAATGATATTAATGCATAC       | 600     |
| Db | 1203926 | AACCATGTAGCTTCTTCAAGAGTAAATCACTTCACTTTTACCATTAATATATGCTT        | 1203867 |
| Qy | 601     | ATCACCATCATCAATCAATAAATAAAAATCTGC-CGTTTGGTGGCTAAGTAAGCTATTTGTG  | 659     |
| Db | 1203866 | GCCATTTTCAGTTTTAGCAATATATCCGCACTATTTTCCGAAAAGAAAGCATATGTTT      | 1203807 |
| Qy | 660     | TGATPATGTGTGATTTGGGGCAAGCTTGTTCAGTGTGTCAAGCATTCCTTGCTAAAT       | 719     |
| Db | 1203806 | CGCTGATTTTTCGTTTTTCCCAAAAGTCAGCC-----                           | 1203775 |
| Qy | 720     | TGGCATGCTTTGGTGAATGATAGGGGTTTGGCAATTTTCAAAAATAGCAATTTTGTG       | 779     |
| Db | 1203774 | ---TATCTTCATTTGTCAACACCTACTTAGCAATATGATAAATATGCTATTTTGCTT       | 1203718 |
| Qy | 780     | CCAAATATATCCACCATGCCATCGTCCGATCAAGATGATCAAGCGAGATTTAAATATG      | 839     |
| Db | 1203717 | GCGATATATCTTTAAATATCCATATAGCGATCCATATGATCTTCACTACGTTCAAGACAG    | 1203658 |
| Qy | 840     | TTGCCCCCTTGAAGCTGCCAAATTTGAGATATGCTCAAGCTGAATACTTGACAGCTTCAGCA  | 899     |
| Db | 1203657 | TCGCTGCGCGAGCTTTTAAAGCATATAGTATAGTGTCTCAAGCTGAATAAAGCTTAACTA    | 1203598 |
| Qy | 900     | CCACCAATATCCATATTTTGGATTCGTTAATATCAAGTCAGAGCGTGCATATTTACCGC     | 959     |
| Db | 1203597 | CATTAAGTTCACAACTTTCATTCC--AACATATGACAAAGCGGGAAATCCCAATATTTCCGC  | 1203541 |
| Qy | 960     | CGACACCAACATCATATGCTGTCATATTTTGGCATCTGCGCTATCTAATATGCTGAACAGTGC | 1019    |
| Db | 1203540 | CCATACCAACTTTCACACACAGAGCTTTCGCACTTTCATTAACATTAAGTATAGCGTATC    | 1203481 |
| Qy | 1020    | TTTTGGCATTTGAGCGCTGTGATGGCGATGATGTGTGTGTTGTTCTGAGTTAGATCTC      | 1079    |
| Db | 1203480 | TTTTTACATTTGAACCTGTAAATCCCAAAATTGGCTTGTGTGCGTCGGGGCA-----       | 1203429 |
| Qy | 1080    | GTGCTTGAAGTGTATCAATGAACAATTTGAACATGCTGTATCAACGAATGCTTGAGCTT     | 1139    |
| Db | 1203428 | -----GAATATTCAAATATGCGCATTAATCTTCCACTCCCGCTTTAA                 | 1203388 |
| Qy | 1140    | TGGCGGCGACATATCTGGCGCTTCTGGGTTAATGCGGGGCTGATGATGATGATGCG        | 1199    |
| Db | 1203387 | GTCGGGTTTGAATTTCTGTGTGTTTTTACCGCAAGCCCTGGGGTAAATTAACAATCATATGCG | 1203328 |
| Qy | 1200    | CTGAGCTTAAAGTTCATCTGTCCAGCTGCCAAATGAACAATGAACGCTGTGGGCACTT      | 1259    |
| Db | 1203327 | TTTCAAGTAAACATTCCTGATTTTAAATCAACAGTATGAAGAAGGATATTTTGAAGGAATT   | 1203268 |
| Qy | 1260    | TATCTGCCAAGTATGAGATTTGGGATTAACATGATGATGCTTAACTTTGTGACCGTGATTTGA | 1319    |
| Db | 1203267 | TATCATATCCAGTATGAGATTTTTCAGATATCAATCAACGAATATTTACCTGTGGGATA     | 1203208 |
| Qy | 1320    | CCAAAAAATTCACAGCTGACGTCAGAGCTGCCAAACGAGACAGCGCATATTTTTT         | 1377    |
| Db | 1203207 | AGAGTATATCCACACAAAGAAAGCTGTTTTTGGCAAGCCGCAATGATATATATTTTTT      | 1203150 |

```

RESULT 9
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: P3186P2CID1
; CURRENT APPLICATION NUMBER: US/10/158.865

```

```

CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 09/557,884
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 08/476,102
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1030121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4416)..(4416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g

```

```

1  SEQ ID NO 1
2  LENGTH: 1830121
3  TYPE: DNA
4  ORGANISM: Haemophilus influenzae
5  FEATURE:
6  NAME/KEY: misc_feature
7  LOCATION: (4747)..(4747)
8  OTHER INFORMATION: n equals a,t,c, or g
9  FEATURE:
10 NAME/KEY: misc_feature
11 LOCATION: (9921)..(9921)
12 OTHER INFORMATION: n equals a,t,c, or g
13 FEATURE:
14 NAME/KEY: misc_feature
15 LOCATION: (10150)..(10150)
16 OTHER INFORMATION: n equals a,t,c, or g
17 FEATURE:
18 NAME/KEY: misc_feature
19 LOCATION: (29298)..(29298)
20 OTHER INFORMATION: n equals a,t,c, or g
21 FEATURE:
22 NAME/KEY: misc_feature
23 LOCATION: (35543)..(35543)
24 OTHER INFORMATION: n equals a,t,c, or g
25 FEATURE:
26 NAME/KEY: misc_feature
27 LOCATION: (36551)..(36551)
28 OTHER INFORMATION: n equals a,t,c, or g
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: (36636)..(36636)
32 OTHER INFORMATION: n equals a,t,c, or g
33 FEATURE:
34 NAME/KEY: misc_feature
35 LOCATION: (40808)..(40810)
36 OTHER INFORMATION: n equals a,t,c, or g
37 FEATURE:
38 NAME/KEY: misc_feature
39 LOCATION: (4416)..(4416)
40 OTHER INFORMATION: n equals a,t,c, or g
41 FEATURE:
42 NAME/KEY: misc_feature
43 LOCATION: (44905)..(44905)
44 OTHER INFORMATION: n equals a,t,c, or g
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: (44975)..(44975)
48 OTHER INFORMATION: n equals a,t,c, or g
49 FEATURE:
50 NAME/KEY: misc_feature
51 LOCATION: (45593)..(45593)
52 OTHER INFORMATION: n equals a,t,c, or g
53 FEATURE:
54 NAME/KEY: misc_feature
55 LOCATION: (45732)..(45732)
56 OTHER INFORMATION: n equals a,t,c, or g
57 FEATURE:
58 NAME/KEY: misc_feature
59 LOCATION: (47036)..(47036)
60 OTHER INFORMATION: n equals a,t,c, or g
61 FEATURE:
62 NAME/KEY: misc_feature
63 LOCATION: (51334)..(51334)
64 OTHER INFORMATION: n equals a,t,c, or g

```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

Query Match 11.6%; Score 160.4; DB 8; Length 1830121;  
Best Local Similarity 49.6%; Pred. No. 9.3e-32;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
QY 1 TAAAGCGTGAAACATGTTAAACAACCTATATCTGCTGCTCATTAATTTTAAATGATC 60
DB 1204439 TAAACATTTGAGCTTAATGCTTAATTTCTGCGCGCTTTTCAAAAGCAACGATC 1204380
QY 61 CATGCTAGCAACAGCGGTGACAAAGCACAGCCCTAGCGGTCTTGACTGCTGCCAA 120
DB 1204379 GAGACTTGACAAAGCGGCGCAATATATCA-----TATCTCGGCTTGCAA 1204332
QY 121 TACTGCCCATCAAGTGACCAACCTGATATATTTAGATGCAAAATGTCACCTT 180
DB 1204331 TGTGGGCGGTAAATTTTCAAGCGCACCATCTGCAATTAATTAATGTTGG 1204278
QY 181 GAGTAAATCTTGCTCAATTAATGCGGCGATCTTGACCGATTAAATACACTATCCACA 240
DB 1204277 CGATGAAATTTTGAAGCGGCGCACCATCTGCAATTAATTAATGTTGG 1204218
QY 241 CTGCATTAACATATGCTGACGCTCACTAATATCTTGACCTTACCAAGCCCAACCAAT 300
DB 1204217 TTGATTAATTAATTCGCTTAATCTGAAATATGAGCCCTTTTCCGCTTCGCTGCAA 1204158
QY 301 CAATGCCAGTATGATATGCAATTAACCGTGCCAAAGCCCATCAATTCGCTGCAATGTTGA 360
```

Db 1204157 CAATGCAATTAACCTCAATATA-----AAGCCAGCCAAATGCAACATCTGACT 1204107  
Qy 361 GCCAGATGTTGATCCTTTGAATCATTAATAATAGCTTGCTCATATCATCAATATA 420  
Db 1204106 CCCACATTTTGTGCTTTAAGTCATTAATCCAAAGAAATGCAATAGCTTGAATGCACTAA 1204047  
Qy 421 TTCAACAGCGATGGGATGAGCTTTGAAAGTCTTAAGGATTAAGCAATGCGATTCATGGG 480  
Db 1204046 TTGAAAACGATGATCTTAACCTTTGAAATGACGAAGTGGCGATGCAATCTAAAT 1203987  
Qy 481 AAGATTGGGAGCTGTGCTTAAGCAAGTGCAGATTAAGGCAATGATAGTATGCTTGGCC 540  
Db 1203986 AATACCTAATAGCTTGCGCAATGCTGTGTGCGCAAAATGTCATATAATTAATGCGACC 1203927  
Qy 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 1203926 AACCAATGATGCTTCTTCAAGATTAATCACTTCTTAACTTTAAATTAATGCTT 1203867  
Qy 601 ATCAACATCATCAATTAATAAATATGTC-CGTTGGTGGTAAAGTAACTTAATTTGTTG 659  
Db 1203866 GCCATTTTCAAGTTTAAAGCAATATCCGCACTATTTTCGCAAAAGAAACGGTATGTT 1203807  
Qy 660 TGAATATGATGATTTGGGGCAACGCTTGTCACTGTGTCAAGCATGCTTGGCTAAAT 719  
Db 1203806 CGCTTGATTTTCTTTTCCCAAAAGTCAGCC----- 1203775  
Qy 720 TGGCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779  
Db 1203774 ---TATCTTATGTTTCAACACACTACTTATGATTAATTAATTAATGCTT 1203718  
Qy 780 CCAATTAATCACAATGCAATGCTGCGGATCAAGATGATGATGATGATGATGATGATGATG 839  
Db 1203717 GCGGATTAATCTTAAATCCATTAAGGATCAATGATCTTCAATCACTTCAAGCAG 1203658  
Qy 840 TTGCCCCCTGAGCTGCAATTTGAGATATGCTCAAGCTGAAATCTTAAGCTCAAGCA 899  
Db 1203657 TCGCTGCCGAGCTTTTAAAGCTTAAGTGTCTCAAGCTTAAATCTAAGAAAGCTTAAGTA 1203598  
Qy 900 CCACCAATTCATATTTTGGATCGTTAATTAATCAAGTGCAGGCGTGCATATTTACCGC 959  
Db 1203597 CATTAAGTTTCAATCTTCACTTC---AACAATGACAAAGGCGAAATCCCAATTTTCGCG 1203541  
Qy 960 CGACACCAACATCATGCTGCAATGTTTGGCATCTGCGCTACTAATATGCTGAACAGTGC 1019  
Db 1203540 CCATACCAACTTTCACACCGCAGCTTTCGCCATTTCAATAACTAAAGTATGATCGTAC 1203481  
Qy 1020 TTTTGGCATTTGAGCGTGAATGAGGATGATGATGATGATGATGATGATGATGATGATG 1079  
Db 1203480 TTTTACCATTTGAACCTGTAATCCCAATTTGGCTTGTGCTGCTGATGATGATGATG 1203429  
Qy 1080 GTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139  
Db 1203428 -----GAATTAATCAATATCGCGATTAATCTTCACTTCGCGCTTAA 1203388  
Qy 1140 TGGCGGCGACATATCTCGCGCTTCTTGGGTTAAATGCGGCGCTGATGATGATGATGATG 1199  
Db 1203387 GTGCGGTTGAATTTCTGGGTTTAAACGCAAGCCCTGGGCTAATTAACAATCATATCCG 1203328  
Qy 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259  
Db 1203327 TTTCAAGTAACTTCTGATTTTAAATCACTAATGAAGAGGATATTTTGAAGAGTT 1203268  
Qy 1260 TATCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319  
Db 1203267 TATCAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1203208  
Qy 1320 CCAAAAAATTCACAGCTGACAGTCCAGAGTGGCCAAACCGACGACGCAATATTTT 1377  
Db 1203207 AAGATTAATCAACAAAGAAAGACTGTTTGGCAAGCCGATGATGATGATGATTTT 1203150

RESULT 10  
US-10-981-687-1/c

Sequence 1, Application US/10981687  
Publication No. US20050131222A1  
GENERAL INFORMATION:  
APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: Nucleic Acid Sequence of the Haemophilus Influenzae Rd Genome, Fra-  
FILE REFERENCE: PB16P2C1D12  
CURRENT APPLICATION NUMBER: US/10/981,687  
CURRENT FILING DATE: 2004-11-05  
PRIORITY FILING DATE: 2002-06-03  
PRIORITY FILING DATE: 2000-04-25  
PRIORITY FILING DATE: 1995-06-07  
PRIORITY FILING DATE: 1995-04-21  
PRIORITY FILING DATE: 1995-04-21  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1630121  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4747)..(4747)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (3921)..(3921)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (40808)..(40808)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44416)..(44416)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44975)..(44975)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45593)..(45593)  
OTHER INFORMATION: n equals a,t,c, or g  
FEATURE:  
NAME/KEY: misc\_feature

```
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
```

```
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
```

Query Match 11.6%; Score 160.4; DB 9; Length 1830121;  
Best Local Similarity 49.6%; Pred. No. 9.3e-32;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
QY 1 TAAAGCTGAACATGTTAAACAACCTATCCGCTCATTTAATTTTAAATGATC 60
DB 1204439 TAAACATTAAGCTTAATTAATTTCTTGGCCGCTTTCAAAAGACAACTGATC 1204380
QY 61 CATGCTAGCACAAGCAGGTGACAAAGCAGCCCTTAGCGGTGACTGCGTGC 120
DB 1204379 GAGACTTGACAGAGCGGAGCAATATACCA-----TATCTCGCTTGCA 1204332
QY 121 TACTGCCCATCAAGGTACCAACCTGATATATTTAGATGCCAAATGTCATCCTT 180
DB 1204331 TGTGGGCGTAAATAATTCATTCCTTTCATTTGATGAAACA-----ATACCTTG 1204278
QY 181 GAGTAATCTGCTCAATTAAGGGGATCTTGACGATTAATACACTATCCACCA 240
```

D 1204277 CGATGAAATTTTGCAAGCAGCGCACTTCGACCAAAACAATAAATATGTGTG 1204218  
Q 241 CTGATAACATATGCTGACAGCTCACTAAATCTTGACCTTACCAAGCCACCAAAAT 300  
D 1204217 TTGATTAATTAATTCAGCTAATTCGAAAATCAAGCCCTTTCCGTCCGCTAGCA 1204158  
Q 301 CAATGCCAGATATATGCAATTAACCGTCCCAAGCCCAATTAATGCTGCATGTTGA 360  
D 1204157 CAATGCAATTTTACCCTCAATTA-----AAGCCCAAGCAATGACGACCTGACT 1204107  
Q 361 GCCGATGTTTGTACCTTTTGAATTAATAATGCTGTATCATATATATCAATTA 420  
D 1204106 CCCCACTTTGCTGCTTAAAGTCAATTAATCAAGATGCTTATGCTGATCACTPA 1204047  
Q 421 TTCAAGCGATGCGTGAACCTTTTGAAGCTTAAGGATCAAGATGCAATGCAATGAGG 480  
D 1204046 TTGAAAACGATGATCTAACCTTTGAATATGCAAGAGCGGTAGAAATTAATTAAT 1203987  
Q 481 AAGATTGCGAGCTGTGCTTAAGCAAGTGCAGATAGGCAATGAGTATGCTTGGCC 540  
D 1203986 AATACCTATGCTGTGCGCAATGCTGTGCGCAAAATGTTCAATATATGCGACC 1203927  
Q 541 TTGATTTTAATGATGATGATGATAAAGTTGTTTTTGAATATATATGCAATACC 600  
D 1203926 AACCAATGATGCTTCTTCAAGATTAATACTTCACTTTTCAATTAATATGCTT 1203867  
Q 601 ATCAACATCATCAATTAATAATAATCTGC-CGTTTGCTGCTAAGTATGCTATGTTG 659  
D 1203866 GCCATTTTGAATTTTGAAGCAATATCCGCACTATTTTCCGCAAAABAAACGATATTT 1203807  
Q 660 TGAATATGCTGTATTTGGGCAACGCTTGTCACTGTGCTCAGATGCTGCTGCTAAT 719  
D 1203806 CGGTTGATTTTCTGTTTTCCCAAAAGTCAGCC----- 1203775  
Q 720 TGGCATGCTGTTGGATGATGAGCGGTTTGGCAATTTTCAAAATACGCAATTTTGTG 779  
D 1203774 ---TATCTTATTTTCAACAACACTTATGATATATATTAATGCTAATTTTGTCT 1203718  
Q 780 CCAATATATCACCATGCGATCGTGCAGATCAAGATGATCAGCGAGAGATTTTAAATG 839  
D 1203717 GGGATATATCTTAAATCAATATAGGATTCATATGATCTTCACTACCTTCAAGCAG 1203658  
Q 840 TTGCCCCGTGAGCTGCCAAATTTGAGATATGCTCAGCTGAAAACCTTGACAGCTCAAGCA 899  
D 1203657 TCGCTCCGCGAGCTTTTAACTTAAGTGTCTCAAGCTGAAAACCTGAAAGCTTAGTA 1203598  
Q 900 CCAACCAATCATATTTTGGATGCTTAATTAATCAAGTGAAGCGTGCCTCAATTTACCCG 959  
D 1203597 CATTAAGTTCAAACTTCTCAATC---AACAATGACAAAGCGGAATCCCAATATTTCCGC 1203541  
Q 960 CGACACCAACATCATGCTGATGTTTGGCATCTGCTACTTAATGCTGAACAGTGC 1019  
D 1203540 CCATACCAACTTTCACACGAGAGCTTGGCCATTTCAATACTTAAGTATGTTACGTTAC 1203481  
Q 1020 TTTTGGCATTTGAGCTGTGATGCGAATGATGCTGTGTTGTTGCTTGAATGATCTC 1079  
D 1203480 TTTTACCATTTGAACCTGTATCCCAATGCTTTTGTGCTGCGCGGCA----- 1203429  
Q 1080 GTGCTTTGAGTATCAATGAACAATGAACATGCTGATCAACAGAAATGCTTGAGCTT 1139  
D 1203428 -----GAATAATTCATATATGCGCAATTACTTCCACTCCGCTTTTAA 1203388  
Q 1140 TGGCGGCACAATATCTGCGCTTCTGGGTTAATGCGGGGCTGATGATGATCTGATCGG 1199  
D 1203387 GTGCGGTTTGAATTTCTGGGTTTAAACGCAAGCCCTGGGCTTAATTAATATATTCGC 1203328  
Q 1200 CTGAGCTTAAGATTCATGCTCAAGCTGCAAAATGACAAATGACAGCTGCTGGGAGATT 1259  
D 1203327 TTTCAAGTAACATTTCTGATTTAAACTACAGTATGAGAGGATATTTTGAAGAAATT 1203268  
Q 1260 TATCTGCCAAGAGATGAGGATTAACATGATGATGATGATGATGATGATGATGATGATG 1319  
D 1203267 TATCAATACCAAGTATTTTTCGAGTATCAATCAACGAATATTAAGCTGTGTTGGATA 1203208

Q 1320 CAAAAAATTCAGAGCTGACAGTCCAGAGCTGCCAAACCGAGAGGCAATATTTT 1377  
D 1203207 AGAGATATATCCACAAAGAAAGACCTGTTTGGCAAGCCGAGATGATATATTTT 1203150  
RESULT 11  
US-10-795-159-569  
; Sequence 569, Application US/10795159  
; Publication No. US20050221439A1  
; GENERAL INFORMATION:  
; APPLICANT: BAXALETZ et al.  
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE  
; FILE REFERENCE: 28335/38815A  
; CURRENT APPLICATION NUMBER: US/10/795,159  
; PRIOR FILING DATE: 2004-03-05  
; PRIOR APPLICATION NUMBER: US 60/453,134  
; NUMBER OF SEQ ID NOS: 771  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 569  
; LENGTH: 14324  
; TYPE: DNA  
; ORGANISM: H. influenzae  
US-10-795-159-569  
Query Match 11.2%; Score 154.4; DB 9; Length 14324;  
Best Local Similarity 49.0%; Pred. No. 2.7e-31;  
Matches 591; Conservative 0; Mismatches 546; Indels 69; Gaps 4;  
Q 172 ATCACTTGAAGTAAATCTTGTCAATTAATCGGGGATCTTGAACGATTAATATACACT 231  
D 5280 ATAGCTTTGATGAAATTTTGGCAAGCTGCACACATCTCGACCAAAACATTAACAAAT 5339  
Q 232 ATCAACCACTGATTAATATATGCTCAAGCTCACTAAATCTTGAACCAAGCC 291  
D 5340 AATGTGTGTTAATTAATTAATCACTAATCTGAAAATCAGCCCTTTTCCGTCC 5399  
Q 292 ACCCAATATCAATGCCAGTATGATATGCCATTAACCGTGCAGCCATGCAATGCTGC 351  
D 5400 GCTTACCAACAAATGCAAT-----TTACCTTAATTAATGAAGCCAGCTAATGCAAC 5450  
Q 352 AATGTTGAGCGGATGTTTGTACTTTGATATCAATTAATATGCTGCTATCAATATC 411  
D 5451 AACTGATACCCCAACTTTGTTGCTTAAAGTATTAATCAAGAAATCACTATGCTTG 5510  
Q 412 ATCAATATATTCACAGCGATGCGGATGAGCTTTGAAAAGTCTTAAGGATATCAAGATGC 471  
D 5511 ATGCACTAATGAAAACGATGATCTTAACCTTTGAAAATGACGAAGTGCAGATTTGA 5570  
Q 472 ATCCATGGGAAGTTGGCAGCTGTGCTTAAGCAATGCAATGCAATGAGTATGATGTT 531  
D 5571 ATCTAATTAATTAATCCAGAGCTGTGCAATGCTGTGCTGCCAAAATGTTCAATATAT 5630  
Q 532 ATGCTTGCCTTGAATTTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 591  
D 5631 ATGCGACCGAGATTAATGATCTTCAACAGCAAAATTAATCTTATCTTTTACCTTAA 5690  
Q 592 TGGCATACATCAACATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 651  
D 5691 GATATGCTTGCATTTTCAATTTTAAAGCAATTAATCCGACATATTTTCCGAAAAGAAAC 5750  
Q 652 ATTTGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711  
D 5751 GGTGTGTCTGCTGATTTTCACTTACCAAAAGTCAAGC----- 5790  
Q 712 GGTAAATTTGAGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 771  
D 5791 -----TATCTCATTTTCAAAATCTCGACCTCAGATTAATGATTAATGCTTAA 5840  
Q 772 TTTTGTGCCAATTAATCAACATGCAATGCTGCGGATCAAGATGATCAAGCGAGAGATT 831  
D 5841 TTTTGTGCGGATATCTTCTAATTCATATAGGATCAATATGATCTTCACTGATGCTG 5900

```
QY 832 TAAATTTGTTCCCTTGAGCTGCGCAATTTGAGATATGCTCAAGCTGAAAACCTTGACG 891
D 5901 CAAAACAGTCGACGCCGCGAGCTTTTAAAGCTATAGCTGTCTCAAGCTGAAAACCTGAAAAG 5960
QY 892 CTCAAGACCAACCAATTCATATTTTGGATCGTTAATTAATTCAGTCAGCGCGCTCAAT 951
D 5961 CTCTAGATACATTAAGTTTCAATCTTCATTC--ACAAATGACAAAGCGGGAATCCCAAT 6017
QY 952 ATTACCGCGCACCAACCAATCATGCTGCAATGTTTGGCATCTCGCTCAATATGTCGT 1011
D 6018 ATTCCGCCCAATCAACCTTTCAGCGCAGCTTTCGCCAATTCATTAATTAAGTAGT 6077
QY 1012 AACAGTCTTTGGCATTTTGAGCTGATGCGATGATGATGATGATGATGATGATGATGAT 1071
D 6078 AACAGTCTTTTGCAATTTTGAACCTGTAATCCCAATTTGCTTGTGCTGCGC--- 6134
QY 1072 TAGATCTGCTGCTTGAAGTGTATCAATGAACATTTGATCGCTGATCAACGAATGCG 1131
D 6135 -----GCAATATTAATTCATTAATTCGCGATTAATTCCTTCACTCC 6170
QY 1132 TTGAGCTTTGGCGCGACCAATATCTGCGCTTCTTGGGTATATGCGCGCGCTGATGAT 1191
D 6171 CGCTTAAAGTGGCTTTGAATTTCTGCTTTTAACTGCGCAAGCCCTGCTAATTAACAT 6230
QY 1192 CTGATCGCTGAGCTTGAAGTTCACTGTCGAAGCTGCAAAATGACAAATGAAAGCGCTGT 1251
D 6231 CATTCGCTTTCAAGTAAACATTTCTGATTTAACTCCAGTATGAAGAGGATATTTTG 6290
QY 1252 GGGAGATTTATCTGCGCAAGTATGAGTTGAGTATCAATGATGATGATGATGATGATGAT 1311
D 6291 AGGAATTTATCAATACCAATAGGCTTTTTCGGGTATCAATCAAGAAATTAAGCTG 6350
QY 1312 GTGATTTGACCAAAATTTTCAAGCTGACAGTCCAGAGCTGCCAAACGACGACGATA 1371
D 6351 TTGGATTAAGATTAATTCACAAAGAAAGCTGTTTTCAGAGCCCAATGATGATAT 6410
QY 1372 TTTT 1377
D 6411 ATTTT 6416

RESULT 12
US-10-795-159-685/c
; Sequence 685, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BARALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 685
; LENGTH: 908766
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (39640)..(39640)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (292404)..(292404)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (292495)..(292495)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (908604)..(908641)
; OTHER INFORMATION: n is a, c, g, or t
US-10-795-159-685

Query Match 11.2%; Score 154.4; DB 9; Length 908766;
Best Local Similarity 49.0%; Pred. No. 2.9e-30;
Matches 591; Conservative 0; Mismatches 546; Indels 69; Gaps 4;

QY 172 ATGACCTTTGAGTAATCTTGCATATTAATGCGGCACTCTTGAACCATTAATTAACACT 231
D 107917 ATAGCTTTGAGTAATTAATTTTGAAGCTGACACCACTTCGACCAAAATTAACAAAT 107858
QY 232 ATCCACCACTGATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCC 291
D 107857 AATGTGTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 107798
QY 292 ACCCAAAATCAATGCAATGATATATGCAATTAACCGTCCAGCCCAATTAATGCTGC 351
D 107797 GCCTAGCACAAAATGCAAT-----TTACCTCAATTAATTAAGCCCAAGCTAATGACG 107747
QY 352 AATGTTGAGCGCATGTTTGTACCTTTGATTAATTAATTAATTAATTAATTAATTAATTA 411
D 107746 AACTGTACTCCCAATTTGTTGTTAGTCAATTAATTAATTAATTAATTAATTAATTA 107687
QY 412 ATCAATATATTCACAGCATGCGGTAGCCCTTTGAAAGTCTTAAAGGTATCAAGCATGCG 471
D 107686 ATGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 107627
QY 472 ATCCATGGAAGATGCGAGCTGCTTAAGGCAATGCAATGCAATTAATTAATTAATTAAT 531
D 107626 ATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 107567
QY 532 ATGCTGCTTGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 591
D 107566 ATGGCAACGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 107507
QY 592 TGCCATACATCACTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 651
D 107506 GTATTCCTGCTGCTTTTCAATTTTGAAGCAATTAATTAATTAATTAATTAATTAATTA 107447
QY 652 ATTTGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 711
D 107446 GGTTCCTGCTGCTTTTCAATTTTGAAGCAATTAATTAATTAATTAATTAATTAATTA 107407
QY 712 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771
D 107406 -----TTATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 107357
QY 772 TTTTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 831
D 107356 TTTTGTGCTGCTTATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 107297
QY 832 TAAATTTGTCCTTGAAGCTGCGCAATTTGATGATGATGATGATGATGATGATGATGAT 891
D 107296 CAAAACAGTCGACGCCGAGCTTTTAAAGCTATTAAGTTGTCTCAAGCTGAAAACCTGAAAAG 107237
QY 892 CTCAAGACCAACCAATTCATATTTTGGATCGTTAATTAATTAATTAATTAATTAATTAATTA 951
D 107236 CTCTAGATACATTAAGTTTCAATCTTCATTC--AACAAATGACAAAGCGGGAATCCCAAT 107180
QY 952 ATTACCGCGCACCAACCAATCATGCTGCAATGTTTGGCATCTCGCTCAATATGTCGT 1011
D 107179 ATTCCGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 107120
QY 1012 AACAGTCTTTGGCATTTTGAGCTGATGCGATGATGATGATGATGATGATGATGATGAT 1071
D 107119 AACAGTCTTTTTCATTTGAACCTGTAATCCCAATTTGCGCTTGTGCTGCGCG--- 107063
```

QY 1072 TAGATCTGCTGCTTGTGAGTATCATGAACAATTGAACTCCGTATCA CAGGAATGCC 1131  
DB 107062 -----GCAGAAATAATTCAATATCGCCGATTACTTCCACTCC 107027  
QY 1132 TTGAGCTTTGGGGGCGCAATACTCGGCGTTCTTGAGTTATGCGGGGCTGATGAT 1191  
DB 107026 CGCTTTAAGTGGCGTTGAAATTTCTGATGTTTTTACCGCAAGCCCGGCTAATAACAAT 106967  
QY 1192 CTGATCGGCTGAGCTTTAAGAGTTCACTGTCGAAGCTCCAAAATGACAAATGAACGCGCTG 1251  
DB 106966 CATATGCGCTTTCAAGTACCATTCCTGATTTAAACTCAAGTATGAAGAGGATATTTTG 106907  
QY 1252 GGGCAGCTTTATCTGCCAAGGTAGGATTTGGAATTCATCATGATATGCTAACTTTGACC 1311  
DB 106906 AGGAAGTTTATCATATCCAGTACGCTTTTTCGGGATATCATCAACGAATATTAGCGCTG 106847  
QY 1312 GTGATTGACCAAAAATTTACAGCTGACATTCAGAGCTGCCAAACCGACGCGATA 1371  
DB 106846 TTGGGATTAAGAGATTAATCCACAAAGAAAGACCTGTTTCCAAAGCCCAATGATCGTAAT 106787  
QY 1372 TTTTTT 1377  
DB 106786 ATTTTT 106781

## RESULT 13

US-10-282-122A-25165/c  
; Sequence 25165, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:

APPLICANT: Wang, Lianguo  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Foreyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITPA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 25165  
LENGTH: 1281  
TYPE: DNA  
ORGANISM: Legionella pneumophila

## US-10-282-122A-25165

Query Match 11.1%; Score 152.8; DB 7; Length 1281;  
Best Local Similarity 51.6%; Pred. No. 1.9e-11;  
Matches 464; Conservative 0; Mismatches 412; Indels 24; Gaps 4;

QY 170 GCATCACTTTGAGTAAATCTTGCTCAATTAATGCGGCACTTGAACGATTAATAACACA 229  
DB 1163 GGAATTAATCTTTGGCTTAAGCCGATTCATCTTGTCAGCATCTTCAACCAATCAAAACAA 1104  
QY 230 CTATCAACCCACCTGCAATATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGC 289  
DB 1103 GAACGAACAATTTCTGAAATCTGCGCTGCTAAATCTTGGAATAATCCGCCCTTTCCCTG 1044  
QY 290 CCACCCAAATCAATGCAATGCAATGATATGATGCCATAACGGTGCAAGCCCAATTAATGCT 349  
DB 1043 CCACCCGCAATTAAGAAATTTTCCCTTGATTAAGACC--TCTATTTCCATTAATATGA 987  
QY 350 GCAATGATGAGCCGATGTTTGTATCTTTGATCAATTAATAATATGCTTGTATCAATA 409  
DB 986 GAGATTTGACACCAATATTTGTTCTTGCATCAATTAATCAACTCACTCATCTCATCTACT 927  
QY 410 TCAATCAATATATCAAGGATGCGGTAGCCCTTTGAAAGTCTTAAGGATCAACATG 469  
DB 926 TCTCTTACCACCTGCGATGATGATGACAGTCCAGGGAAGTTTAAACATTAAGAAATG 867  
QY 470 GCATCCATGGGAAATTTGGCAGCTGCTGCTAAGGCAAGTCAAGATTAAGGATTTAGTAG 529  
DB 866 TGTGATTAAGAAATTTCCCGCTTGGCTTAAGCAAGCAAGCAAGCCATTCATCCAA 807  
QY 530 TTAATGCTTCCCTTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 589  
DB 806 TTAATGACCCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 747  
QY 590 AATGCAATACCATCAATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 649  
DB 746 AATTAAGTATATCTCTTCAATCAAGCCCAATTTCCCATTT-----GAAG 698  
QY 650 CTATTTGTTGTATTAATGATGATTTGGGGCAAGCTTGCAGTGTGTCAGCAATGAC 709  
DB 697 GCGCGCTTTGGCAAAAGATGATTTTAAATCTGCTTGAACGATTTGTGAGCA---- 642  
QY 710 TTGGCTAAATTTGCAATGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 769  
DB 641 ---ACAGTATACGCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 585  
QY 770 AATTTTGGCCAAATTAATCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 829  
DB 584 TGTTTTGTGCTGATTAAGGCTTTCATAGTATGCTGCTTCCAAATGATCAGGCGTTTACA 525  
QY 830 TTTAAATATGTTGCCCTTGAAGCTGCAATTTAGATATGCTCAAGCTGAATCTTGAC 889  
DB 524 TTAAGAAATGTCGCAACCACTGCGCAAGAAATTAAGTTAAATCAATTAATTAATTAAT 465  
QY 890 AGCTCAAGACCAACAATC---CATATTTGATGCTTAATTAATTAATTAATTAATTAATTA 946  
DB 464 AACTCAATATCAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405  
QY 947 CCAATTTTCCGCGCAACCAATCAATGCTGATGATGATGATGATGATGATGATGATGATGAT 1006  
DB 404 CCGATATTTTCCGCTAAGCTTACGGAACCCGCGCGCTTGAACCAATTTCTCCACCAAA 345  
QY 1007 GTCTTAACAGTCTTTTGGCAATTTGAGCTGATGATGATGATGATGATGATGATGATGAT 1066  
DB 344 GTCTTACAGTGAATTTTACATTAATGCTGCTTAATGCAATCAAGGGGCGCTAATCTCT 285

## RESULT 14

US-10-958-216-402/c  
; Sequence 402, Application US/10958216  
; Publication No. US20050181388A1  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, ALBD



```
APPLICANT: DHARMSI, AKIL
APPLICANT: VEDADI, MASOUD
APPLICANT: ALAM, MUHAMMAD ZAHOR
APPLICANT: ARROWSMITH, CHERYL
APPLICANT: ARREY, DONALD E.
APPLICANT: BEATTIE, BRYAN
APPLICANT: BOZADZIJ, KRISTINA
APPLICANT: CANADIEN, VERONICA
APPLICANT: DOMAGALA, MEGAN
APPLICANT: HOUSTON, SIMON
APPLICANT: KANAGARAJAH, DHUSHY
APPLICANT: LI, QIN
APPLICANT: MANSOURY, KAMRAN
APPLICANT: McDONALD, MERRY-LYNN
APPLICANT: NETHERY, KATHLEEN
APPLICANT: NG, IYV
APPLICANT: OUTANG, HUI
APPLICANT: PINDER, BENJAMIN
APPLICANT: RICHARDS, DAMN
APPLICANT: TAI, MATTHEW
APPLICANT: THAKARADA, ROSANNE
APPLICANT: VALLEE, FRANCOIS
APPLICANT: VIRAG, CRISTINA
FILE REFERENCE: IPT-205_01
TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM BACTERIA
CURRENT APPLICATION NUMBER: US/10/958,216
PRIOR FILING DATE: 2004-10-04
PRIOR APPLICATION NUMBER: PCT/CA03/00462
PRIOR FILING DATE: 2003-04-02
PRIOR APPLICATION NUMBER: 60/369,511
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 60/385,089
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/385,751
PRIOR FILING DATE: 2002-06-04
PRIOR APPLICATION NUMBER: 60/386,553
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,577
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,367
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,566
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,390
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: 60/386,601
PRIOR FILING DATE: 2002-06-06
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1132
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 402
LENGTH: 1314
TYPE: DNA
ORGANISM: Haemophilus influenzae
US-10-958-216-402

Query Match      11.0%; Score 152.4; DB 9; Length 1314;
Best Local Similarity 49.2%; Pred. No. 2.5e-31;
Matches 678; Conservative 0; Mismatches 611; Indels 89; Gaps 7;
```

```
Db 1146 CGATGAAATTTTGCAAGCAGCCGACATCTTCAGCAACCAATTAACAAATATGTGTGG 1087
Qy 241 CTGCATTAACATATGCTGACAGCTCACTAAATCTTGAAGCTTTTACCAAGCCCAACCAAT 300
Db 1086 TTGATTATTAATTAATTCAGCTAATTCGTGAAATAATGAGCCCTTTTCCTCGCTACCA 1027
Qy 301 CAATGCGAGTATGATATGCGATTAACCCGCGCAAGCCCAATCAATGCTCAATGTGTGA 360
Db 1026 CAAATGCAATTTTACCTCAATATA-----AAGCCGCAATGCAACATGTAAT 976
Qy 361 GCCAGTGTGTGACCTTTTGAATCAATTAATAATATGCTTCTATCATATATCATATATA 420
Db 975 CCCACATTTGTGCTTTAGAGTCATTAATCCAAAGATCCATTAAGCTTGAATGACATTA 916
Qy 421 TTCAAGCGATGCGGTGACCTTTTGAAGTCTTAAGGATACAGCATGSCATCATATGAG 480
Db 915 TTGAAAAAGATGATCTTAACCTTTGAAATGACCAAGGCGGTACGAATTAATTAAT 856
Qy 481 AAGATTGCGAGCTGTGCTTAAGCAAGTGAATTAAGCATTTAGATAGTTATGCTTGGC 540
Db 855 AATACCTATAGCTGTGCGCAATGCTGTGCGCAAAATGTCATTAATTAATGCGAGCC 796
Qy 541 TTGATTTTAAATTTGATGATGTGTAAGTTGTTTGTGTTTGTGATATATATGCAATAC 600
Db 795 AACCAATGATGCTTTTCAAGATTAATCACTTATCTTTTACCATTTAAATATGCTT 736
Qy 601 ATCACAATCAATCAATTAATAATAATCTGC-CGTTTGATGCTTAAGTAACTTATTTG 659
Db 735 GCCATTTTCAATTTTAAAGCAATTAATCCGCAATATTTTCCGCAAAAGAAACGATGCT 676
Qy 660 TGAATATGATGATTTTGGGCGCAAGCTGTCTGATGTGTCAGCATTTGCTTGGCTAAT 719
Db 675 CGCTTGAATTTTCGTTTCCCAAAAGTCAGCC----- 644
Qy 720 TGCAATGCTTGTGTGATGATATGCGGTTTGCGCAATTTTCAAAATTAAGCAATTTTGTG 779
Db 643 ---TATCTTATGTTTCAACACACCTTATGACATTAATTAATTAATGCTTATTTGCTT 587
Qy 780 CCAATATATCCACATGCGATGCGGATCAAGATGATCAGCGAGATTAATAATTTG 839
Db 586 GCGATATATCTTAAATATCAATATGAGATTCATATGATCTTCAATGCAAGTCAAGCAG 527
Qy 840 TTGCCCCGTGAGTGCCTTAATTTGAATATGCTCAAGCTGAAATCTTGAAGCTCAAGCA 899
Db 526 TCGCTGCGGAGCTTTTAAGCTTAAGCTGTCTCAAGCTGAAATCTTGAAGCTTATGTA 467
Qy 900 CCACCAATTCATTTTGGATGTTAATTAATTAATCAAGTGAAGGCGTGCCTTAATACCG 959
Db 466 CATTAAGTTCAATATCTCATTC--AGCAATGACAAAGGGAATCCCAATATTTCCG 410
Qy 960 CGACACCAACATATGCTGCTGATGTTTGGCATCTCGCTCACTAATATGCTGAACAGTGC 1019
Db 409 CCAATCAACTTAATCTTAAGCAGCTTGGCCATTTCAATTAATTAAGTATGATACGTTAC 350
Qy 1020 TTTTGGCATTTGAGCTGTGATGCGATGATTTGCTGTGTTGCTTGAATGATCTC 1079
Db 349 TTTTACCAATTAAGTCTGTAATCTCCACAAATGGCTTTGCTGCGCGCGCA----- 238
Qy 1080 GTGCTTTGATGATGAATGAATGAATGAATGAATGCTGATCAAGAAATGCTTGAAGCT 1139
Db 297 -----GATTAATTAATTAATGCGGATTAATTTTCACTCCCGCTTTAA 257
Qy 1140 TGGCGGCAATATCTGCGGTTCTTGGTTAAATGCGGAGCTGATGATGATGATGATGCG 1199
Db 256 GTGCGGTTGAATTTCTGAGTTTAAACGCAAGCCCTGCGGCTAATTAACATATATGCG 197
Qy 1200 CTGAGCTTAAGATTTCACTGTCCAACTGCGCAAAATGAACATGAACGCTGTGCGCACT 1259
Db 196 TTTCAAGTAACTATCTGATTTTAACTACAGTAAAGAGGAGATATCTTGAAGAACTT 137
Qy 1260 TATCTGCCAAGGATGATGGAATTAACATGATGATGATTAATTTGATGACGCTGATTTGA 1319
```

Db 136 TATCAATACAGAGAGATTTTTCGAGTATCAACAGAAATATTAGCCTGTGGATA 77  
 QY 1320 CCAAAATTCAGAGCTGACAGTCCAGAGCTGCCAAACCGACAGGATATTTT 1377  
 Db 76 AGAGATATATCCACAGAAAGACCTGTTTTCGCAAGCCGATGATATATTTT 19  
 RESULT 15  
 US-10-282-122A-30713/C  
 / Sequence 30713, Application US/10282122A  
 / Publication No. US20040029129A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Wang, Liangsu  
 / APPLICANT: Zamudio, Carlos  
 / APPLICANT: Malone, Cheryl  
 / APPLICANT: Heselbeck, Robert  
 / APPLICANT: Ohlsen, Karl  
 / APPLICANT: Zyskind, Judith  
 / APPLICANT: Wall, Daniel  
 / APPLICANT: Trawick, John  
 / APPLICANT: Carr, Grant  
 / APPLICANT: Yamamoto, Robert  
 / APPLICANT: Foreyth, R.  
 / APPLICANT: Xu, H.  
 / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 / FILE REFERENCE: ELITRA, 034A  
 / CURRENT APPLICATION NUMBER: US/10/282,122A  
 / CURRENT FILING DATE: 2003-02-20  
 / PRIOR APPLICATION NUMBER: 60/191,078  
 / PRIOR FILING DATE: 2000-03-21  
 / PRIOR APPLICATION NUMBER: 60/206,848  
 / PRIOR FILING DATE: 2000-05-23  
 / PRIOR APPLICATION NUMBER: 60/207,727  
 / PRIOR FILING DATE: 2000-05-26  
 / PRIOR APPLICATION NUMBER: 60/230,335  
 / PRIOR FILING DATE: 2000-09-06  
 / PRIOR APPLICATION NUMBER: 60/230,347  
 / PRIOR FILING DATE: 2000-09-09  
 / PRIOR APPLICATION NUMBER: 60/242,578  
 / PRIOR FILING DATE: 2000-10-23  
 / PRIOR APPLICATION NUMBER: 60/253,625  
 / PRIOR FILING DATE: 2000-11-27  
 / PRIOR APPLICATION NUMBER: 60/257,931  
 / PRIOR FILING DATE: 2000-12-22  
 / PRIOR APPLICATION NUMBER: 60/267,636  
 / PRIOR FILING DATE: 2001-02-09  
 / PRIOR APPLICATION NUMBER: 60/269,308  
 / PRIOR FILING DATE: 2001-02-16  
 / Remaining Prior Application data removed - See File Wrapper or PALM.  
 / NUMBER OF SEQ ID NOS: 78614  
 / SOFTWARE: Patentin version 3.1  
 / SEQ ID NO 30713  
 / LENGTH: 1305  
 / TYPE: DNA  
 / ORGANISM: Pasteurella multocida  
 / US-10-282-122A-30713  
 Query Match 10.3%; Score 142.2; DB 7; Length 1305;  
 Best Local Similarity 48.8%; Pred. No. 1.7e-28;  
 Matches 577; Conservative 0; Mismatches 528; Indels 78; Gaps 4;  
 QY 195 CAATTAATCGGGGATCTTGAACGATTAATACACTATTCACCCACTGATACATATG 254  
 Db 1126 CAAGTGTTCGATCTTGAACCAATAGCAATTAATATCGGTGATTAATTAAC 1067  
 QY 255 CTGACAGCTCATTAAATCTTGAACCTTTACCAAGCCCAAAATCAATGCGAGTANG 314  
 Db 1066 TGGCTAATTCGAAAGTCCGACCTTTACCATCAACCGAGTAACTAATGCAATGTCC 1007  
 QY 315 ATATGCGATTAACCGTCCCAAGCCATCAATGCTGCAATGTTGAGCGGATGTTGAC 374  
 Db 1006 CGTTTAATCT-----GTAAAGCCAGTTAATGCGGCAACAGTACTGCCCACTTGTGG 956

QY 375 CTTTGAATCATTAATAATATGCTTCTGATATCATATCATTAATATATCAACGATGCG 434  
 Db 955 CTTTGAATCATATCAACCGACGCTACACATGAGCAAAATAGATTAATGGAACGGATGAT 896  
 QY 435 GTAGCCCTTTGAAAGCTTAAAGGTATCAAGCATGSCATTCATGGAAGAATGGACGCTG 494  
 Db 895 CTAGCCCGTTAACTCAACGCAATGACGAGAAATCTCTTAATGAAATACCGGACGCT 836  
 QY 495 TGGCTAAGGCAAGTGCAGATTAAGCATGAGTAGATGCTTATGCTTGAATTTAAAT 554  
 Db 835 GTGCTAATGCGATGAGCGCCGACGATCATTAATAGTAAGCGTCCAACTAATTTATCT 776  
 QY 555 GATGATGCTTAATAAGTTGTTTGTGATATTAATGCAATACCATCAATCATCA 614  
 Db 775 CATCAAGCAACCAACCGTGTCTTGGCCATCAAAATATGACCGCCGTTTGGGTTT 716  
 QY 615 ATTAATTAATAATGCTGCTTGTGCTAAGTAAGTATTTGTTGATTAATGCTGAT 674  
 Db 715 TTAACTAATTAATGATGATGCTTCTCC----- 689  
 QY 675 TTGGGGCAAGCTGTCAGTGTGATGATGATGATGATGATGATGATGATGATGAT 734  
 Db 688 -----CAAGCTTACTCATTTTAAATCCGTCATGCGGTGATGATGATGATGAT 635  
 QY 735 AGATGATGCGGTTTGGCAATTTTCAAAATACGCAATTTTGTGCCAATATCCACCA 794  
 Db 634 TATACACCGCATGTTGGGATGATGATGATGATGATGATGATGATGATGATGAT 575  
 QY 795 TGGCATGCTGCGATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 854  
 Db 574 AGTCACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 515  
 QY 855 CCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914  
 Db 514 TTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 458  
 QY 915 TTTGATGCTTAATAATTAAGTGCAGGCGTCAATATTAATGATGATGATGATGAT 974  
 Db 457 GTGCTGTTCTTAATAATTAAGTGCAGGCGTCAATATTAATGATGATGATGATGAT 398  
 QY 975 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034  
 Db 397 AGCCGCGCATTTTACATGATGATGATGATGATGATGATGATGATGATGATGAT 338  
 QY 1035 CTGTGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1094  
 Db 337 CGGTATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 305  
 QY 1095 CAATGAACATTAATGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1154  
 Db 304 GGCAGAAAGCTCAATATGCGCAATCACTCAATATCTGCTCAAGTCAATGATGAT 245  
 QY 1155 TGGGCTTCTTGGGTTAATGCGGCGGCTGATGATGATGATGATGATGATGATGAT 1214  
 Db 244 CTGTGTTTGAACCGCTATCTTGAATTAATTAATTAATTAATTAATTAATTAATTA 185  
 QY 1215 CACTGCAAGCTGCGCAAAATGACATGAACCGCTGTGGGAGTTTATGCGCAAGGTAG 1274  
 Db 184 GTTGGTTTAAATGCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 125  
 QY 1275 GATGGATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334  
 Db 124 GCTGGCTTCTGATTAATCAACCGCATGCTGTTGTTTCCAGTAGAATCAACAC 65  
 QY 1335 CTGACAGTCAAGCTGCGCAAAACCGACGACCGCATATTTT 1377  
 Db 64 AAGATGAGCTGTGTTTCCCAAAACCAATCAAGTAACTTTT 22  
 Search completed: May 12, 2006, 13:26:10  
 Job time : 1267 secs



Db 2180 GAGACTTGCAAGAGCGGCAATTAATACCA-----TATCTCCGCTTTGCCAA 2133  
Qy 121 TACTGCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAATGTGATCAGCTT 180  
Db 2132 TGTGGGCGTAAATTTCTATCGCTTTTCATTTGTATCGAACAA-----ATATCTTGT 2079  
Qy 181 GAGTAAATCTTGTCTCAATTAATCGGGGATCTTGAACGATTAATATACACTATCCACCA 240  
Db 2078 CGATGAAATTTTGCAGAGCAGCGCATCTCGACCAAAACAATTAATATGTGTG 2019  
Qy 241 CTGCATTAACATATGCTGACAGCTCACAATAATCTTGACCTTTTACAGCCCAACCAAT 300  
Db 2018 TTATTAATTAATTAATGAGCTAATTTCTGAAATAACGCCCTTTTCCGTCCGCTAGCAA 1959  
Qy 301 CAATGCCAGTGAATATGCAATTAACCGTCCAGCCCAATCAATTTCTGCAATGTGTA 360  
Db 1958 CAATGCAATTTACCCCTCAATATA-----AAGCCAGCCCAATGACAGCACTGACT 1908  
Qy 361 GCCGATGTTTGAACCTTTTGAATCAATTAATATGCTGTATCAATATCATCAATATA 420  
Db 1907 CCCCAATTTTGTGCTTTAGATCATTAATCAACGAATGCCATTTAGCTTGAATGACTAA 1848  
Qy 421 TTGACAGGATGGGTAGACCCCTTTGAAAGCTTAAAGGTATCAAGCATGCAATCCATGG 480  
Db 1847 TTGAAAAAGATGATCTTAACCTTTGAAATGACGAAGTGGGTAGAAATTAATTAAT 1788  
Qy 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATTAAGTGTATGCTTGGC 540  
Db 1787 AATACCTAATAGCTTGTGCGCAATGCTGTGTGCGCAAAATTTCAATTAATTAAGGCAAC 1728  
Qy 541 TTTGATTTTAAATGATGATGTGTAAGTTGTTTGTGATGATATATGATCCATACC 600  
Db 1727 AACCAATGATGCTTCTTCAACAGTAAATCACCTTCACTTTTACATTAATATGCTT 1668  
Qy 601 ATCAGCATGATCAATTAATTAATTAATCTGC--CGTTGTGTGCTTAAGTATTTGTG 659  
Db 1667 GCCATTTTCACTTTTACCAATTAATCCGATTAATTTTCGCAAAAGAAACGGTATGTT 1608  
Qy 660 TGAATATGATGATTTGGGGCAACGCTTGTCACTGTGTGCAAGCATTTGGGCTAAT 719  
Db 1607 CGCTTAATTTTCTTTTCCCAAAAGTCAAC----- 1576  
Qy 720 TGGCATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAATTAACGCAATTTTGTG 779  
Db 1575 ---TATCTTCACTTGTTCACACACCTACTTACCATTAATTAATGCGTAAATTTTGCTT 1519  
Qy 780 CCAATTAATCAACCATGCTGCTGCGCATCAAGATGATGAGGGAGAGATTTAAATTTG 839  
Db 1518 GCGGATTAATCTTTTAATCAATATAGGATCAATATGATCTTCAATCAAGTCAAGCAG 1459  
Qy 840 TTGCCCCCTGAGCTGCCAATTTGAGATATGCTCAAGCTGAAATTTGACAGCTCAAGCA 899  
Db 1458 TGGCTGCGGAGCTTTTAAGCTATAAGTGTCTCAAGCTGAAATTAAGAAAGCTCTAAGTA 1399  
Qy 900 CCAACCAATCAATTTTGTGATGCTTAATTAATCAAGTCAAGGCGTCCCAATTTAACCG 959  
Db 1398 CATTAAGTTCACAAATCTTCAATC---AACATGACAAAGGGGAATCCCAATATTTCCG 1342  
Qy 960 CGACACCAACAAATGAGCTGCAATTTTGGCAATCTGCGCTAATATGCTGAACGATG 1019  
Db 1341 CCAATCAAACTTTCACACAGAGCTTTTGGCAATTTCAATTAAGTAAAGTAAAGTAAAG 1282  
Qy 1020 TTTTGGCATTTGAGCTGTGATGCGATGATGCTGTGTGTTGTTGCTTGAAGTAAATCTC 1079  
Db 1281 TTTTACCAATTTGAACTGTAAATCCCAATTTGCTTGTGTGCTGCGCGCA----- 1230  
Qy 1080 GTGCTTTGATGTATGAATGAACAAATGAACATGCTGATCAACAGAAATGCTTTGAGCTT 1139  
Db 1229 -----GAATAATTCATATCGCGATTAATCTTCACTCCGCTTTAA 1189  
Qy 1140 TGGCGGCGCAATPACTCGGCTTCTGCGGTAAATGCGGGGCGTGAATGATGATCTATCG 1199  
Db 1188 GTGCGGTTGAATTTCTGAGTGTAAATCCGCAAGCCCTGGGCTAATTAACAAATCATATCG 1129

Qy 1200 CTGAGCTTAAGAGTTGACTCTCCAGTGCCTCAAAATGACAAATGAACCGCTGTGGCACTT 1259  
Db 1128 TTTCAGATTAACATCTCTGATTTTAACTACAGATTAAGAGGAGATTTTGTGAAGT 1069  
Qy 1260 TATCTGCAAGTGAATGATGATGATTAATCAATGATGATGATTAATCTTGTGACCGCTGAATGA 1319  
Db 1068 TATCAATACAGTAAAGATTTTTCAGATTAATCAACAGAAATTAATGACCTGTGGGATA 1009  
Qy 1320 CCAAAATTTCAACAGCTGACGATGACAGCTGCTCCCAACCGACAGGCAATATTTT 1377  
Db 1008 AGAGATTAATTCACACAAAGAAAGCTGTTTGTGCAAGCCGATGATTAATATTTT 951

## RESULT 2

US-11-194-246-205/c  
; Sequence 205, Application US/11194246  
; Publication No. US20050272089A1  
; GENERAL INFORMATION:  
; APPLICANT: Mott, John  
; APPLICANT: Trepod, Catherine  
; APPLICANT: Atkinson, Stafean  
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME  
; FILE REFERENCE: US1 (Mar 268,05920101)  
; CURRENT APPLICATION NUMBER: US/11/194,246  
; PRIOR FILING DATE: 2005-08-01  
; PRIOR APPLICATION NUMBER: US/10/274,586  
; PRIOR FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: US 60/345,438  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 621  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 205  
; LENGTH: 2850  
; TYPE: DNA  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: Sequence of the mraX coding sequence and flanking regions.  
US-11-194-246-205

Query Match 9.6%; Score 132.8; DB 17; Length 2850;  
Best Local Similarity 54.0%; Pred. No. 5.1e-25;  
Matches 337; Conservative 0; Mismatches 257; Indels 30; Gaps 2;  
Qy 754 ATTTTCAAAATTAACGAATTTTGTGCAAAATTAATCAACATGCTGCGGATCAAG 813  
Db 2839 ATTATGATTAATGCGTAATTTTGTGCGGATTAATCTTAAATCATATAGGATCCAT 2780  
Qy 814 ATGATCAGCGAGAGATTTAAATTTGTGCTGCTGAGCTGCCAATTTGAGATATGCTC 873  
Db 2779 ATGATCTTAAGTCACGTTCAAGACATGCTGCTGCGGAGCTTTTAAAGCTTAAGTGTCTC 2720  
Qy 874 AAGCTGAAATCTTGAACGCTCAAGCAACCAATCAATATTTTGGATGTTAAATATTC 933  
Db 2719 AAGCTGAAATCTTGAAGAGCTCTAGTAATTAAGTTCAACATCTTCAATC---AACATGA 2663  
Qy 934 AAGTGAAGCGCTGCAATTTTACCGCGCACCAACAAATCAATGCTGCTGATGTTTGGCAT 993  
Db 2662 CAAAGGGGAATCCCAATTTTCCGCCCATCAACTTTCACACAGCAGCTTTGCGCAT 2603  
Qy 994 CTGCTACTAATATGCTGAACAGTCTTTTGGCATTTGAGCTGTGATGGGATGTTGG 1053  
Db 2602 TTCAATTAATTAAGATGATTAAGCTTCTTACATTTTGAACCTGTAAATCCCAATATGG 2543  
Qy 1054 TGTGTTGTGCTTGAATGATCTGCTGCTTTGAGTGAATGAATGAATGAATC 1113  
Db 2542 CTTTGTGCTGCGG-----GAGAAATTAATCAATATC 2510  
Qy 1114 GCTGATCAAGAAATGCTTGAAGCTTTTGGCGGCGCAATATCTGCGCTTTGGGTTAAT 1173  
Db 2509 GCCGATTAATCTTCACTCCGCTTTTAAAGTGGGTTTGAATTTCTGAGTGTAAATCCGCAAG 2450

| Query Match           | 9.6%           | Score 132.8  | DB 17     | Length 5640 |
|-----------------------|----------------|--|-----------|-------------|
| Best Local Similarity | 54.0%          | Pred. No. 6,4e-25  |           |             |
| Matches 337           | Conservative 0 | Mismatches 257   | Indels 30 | Gaps 2      |
| QY                    | 754            | ATTTCACAAAATAGCATTTTGTGTGCCAAATATATCCACGATGCATGTCGGCATAG       | 813       |             |
| Db                    | 5626           | ATTATGATTAATGGTAAATTTTGTGTGGCGATATCTTTAAATCCATATAGCATTCAT      | 5567      |             |
| QY                    | 814            | ATGATCAGCGAGATTTAAATTTGTGCCCCCTGAGCTGCCAAATTTGAGATATGCT        | 873       |             |
| Db                    | 5566           | ATGATCTTCAGTACGTTCAAGACAGTCGCTGCGCGAGCTTTTAAAGTATTAAGTGTCTC    | 5507      |             |
| QY                    | 874            | AAGTGAATACTTACACGCTCAGACACCAATTCATATTTTGGATCGTTAATAATTC        | 933       |             |
| Db                    | 5506           | AAGCTGAAAACTAATAAAGCTCTAGTACATAAAGTTCAACATCTTCATTC--AAACATGA   | 5450      |             |
| QY                    | 934            | AAGTCAGGGGTGCGCATATTTACCGCGGACACCAACATCATGTCGTCGATGTTTGGCAT    | 993       |             |
| Db                    | 5449           | CMAAGCGGAAATCCCAATATTTCCGCCCATACCAACTTTCACACACAGAGCTTTCGGCAT   | 5390      |             |
| QY                    | 994            | CTGCGCTACTAATGTCTGAACAGTCTTTTGGCATTTGAGCGCTGTGATGGCGATGATTTGG  | 1053      |             |
| Db                    | 5389           | TTCATAAACTAAGATAGTTACGGTAACTTTTACATTTGAACCTGTATATCCCCACAATTTGG | 5330      |             |
| QY                    | 1054           | TGTGTTTGTGCTTGAAGTTAGATCTGCTGCTTGAAGTATCAATGAACAATTTGAACATC    | 1113      |             |
| Db                    | 5329           | CTTGTGTGCTGCGCGG-----GGAGATATATTCATATATC                       | 5297      |             |

| Query Match           | 5.7%           | Score 78.4  | DB 17    | Length 1323 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 54.2%          | Pred. No. 1.6e-10   |          |             |
| Matches 182           | Conservative 0 | Mismatches 151  | Indels 3 | Gaps 1      |
| Qy                    | 754            | ATTTTCAAAATACGCAATTTTGTGGCCAAATATCCAGCATGCGCATCGATCAAG      | 813      |             |
| Db                    | 666            | ATTTTTCARATATTCACCTTTGCAGAAARATATTCATTCATTTTATATGTAATCCA    | 607      |             |
| Qy                    | 814            | ATGATCAGCGAGAGATTTTAAATTTGTTGCCCTTGAGCTGCCAAATTTGAGATATGTC  | 873      |             |
| Db                    | 606            | ATGATTTATGAAATATTTAGAACATATAGCTATATATGACGAAATAGTGAACATGCTCG | 547      |             |
| Qy                    | 874            | AAAGTGAACCTTGACAGCTCAAGCACCAACCAATCATATTTTGGATCGTTAATA--A   | 930      |             |
| Db                    | 546            | AAAGTGAACCTTGAAAGTTCTTAAACAAACAACTTACTTTTTCACAAACATATA      | 487      |             |
| Qy                    | 931            | TTCAAGTCAGGGGTGCGCATATTTAACGGCGACACCAACATGATGCTGCGATGTTTGC  | 990      |             |
| Db                    | 486            | TTTCGATATGAGAGTTCCAAATATTTACCAACAAACAACTTATACCTTGTTCTACCAA  | 427      |             |
| Qy                    | 991            | CATCTGCGCTTACATATGCTGAACAGAGCTTTTGGCATTTAGAGCTGTGATGCGATAT  | 1050     |             |
| Db                    | 426            | CATTGCGCACTAATTAATGTTATGTTTTCCTCATTTGCTCCCTTATATGCAATGAC    | 367      |             |
| Qy                    | 1051           | TGGTGTGTTTGTGCTTGAAGTGAATCTGCTGCTTT                         | 1086     |             |
| Db                    | 366            | TTTTTCATTTAGCAACTGATACCATGCTGATGTTGTG                       | 331      |             |











```
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: FROM LAMSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10005
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10005
```

```
Query Match          2.9%; Score 40; DB 17; Length 1362;
Best Local Similarity 53.1%; Pred. No. 3.4;
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```
QY 381 AATCAATTAATAATGCTTGTCTATCAATATATATATCAAGCGATGCGGTAGCC 440
    |||||
DB 979 AATCAATTAATAAGATATATCAATCTTTCCCTTATATTAAGATGCTTCTTACCC 920

QY 441 CTTTGAAGCTTTAAGGATATCAAGATGCGATCATCAGGAAGTTGGCGCTGTGCTTA 500
    |||||
DB 919 CTTTGAAGCTTTAAGGATATCAAGATGCGATCATCAGGAAGTTGGCGCTGTGCTTA 860

QY 501 AGCAAGTGCAGATTAAGGATGAGTAGTATGCTTATGCTTGC 540
    |||||
DB 859 CACCTATTGCACTAATGATCATTAAGAAATATATATTTTCCC 820
```

```
RESULT 14
US-10-793-626-4356/c
; Sequence 4356, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P33480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4356
; LENGTH: 3444
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4356
```

```
Query Match          2.9%; Score 39.4; DB 10; Length 3444;
Best Local Similarity 65.2%; Pred. No. 6.7;
Matches 58; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
```

```
QY 5 GCGTGAACCATGTTAACAACCTTATCAGCTGCTCATTTAATTTAATTTGATCCATG 64
    |||||
DB 378 GCTCGGATATCTATGATGAACTTCTCACGAGTTCCTCAATGATGATGATGATCCAA 319

QY 65 CTAGCACAAGCAGGTGACAAAGACAGC 93
    |||||
DB 318 CTGACACAAGCTGTGATTAATAGAACAC 290
```

RESULT 15

```
US-10-750-185-47750
; Sequence 47750, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47750
; LENGTH: 3500
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-47750
```

```
Query Match          2.8%; Score 38.8; DB 10; Length 3500;
Best Local Similarity 51.1%; Pred. No. 9.7;
Matches 91; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
```

```
QY 364 GATGTTGTACCTTTTGAATCATTAATAATATGCTTGTATCATATATCATATATATTC 423
    |||||
DB 1048 GATGTAATCACTTAATTAATCACTTGTAGTAAAGACATTAATCTTCCATTAATGTA 1107

QY 424 ACAGCATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCGATCGAGAG 483
    |||||
DB 1108 ATAGTCTGACCCATCATCATGTAAGCCCTTAAGAGAAATATCTGAGATTCCTCGAGAAA 1167

QY 484 ATTGAGCATGTCCTTAAGGCAAGTCAGATTAAGCATGAGTAGGTTATGCTTGCCT 541
    |||||
DB 1168 GAAGTCTGTAATTAATTTGCTTCAAGTCAAGTCAATTAAGCTTCTTCCCT 1225
```

Search completed: May 12, 2006, 14:02:35  
Job time : 2200 secs

**THIS PAGE BLANK (USPTO)**

STIC-Biotech/ChemLib

190213

From: Chan, Christina  
Sent: Tuesday, May 16, 2006 3:51 PM  
To: Graser, Jennifer; STIC-Biotech/ChemLib  
Subject: RE: rush search

Please rush. Thanks Chris

Chris Chan  
TC 1600 New Hire Training Coordinator and SPE 1644  
(571)-272-0841  
Remsen, 3E89

-----Original Message-----

From: Graser, Jennifer  
Sent: Tuesday, May 16, 2006 2:43 PM  
To: Chan, Christina  
Subject: rush search

Hi Christina,

Will you please authorize the following rush search a special case which is due?

Thanks,  
Jennifer

-----  
STIC:

Please search oligonucleotides of at least 22 nucleotides in length from 11357 to 12736 of SEQ ID NO: 35 from application 10/672787 in pending and commercial databases.

Thanks,  
Jennifer Graser  
Primary Examiner  
Art Unit 1645  
REM 3B09  
2-0858

RECEIVED  
MAY 16 2006  
STIC/GEN. DIV.  
(STIC)

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway;  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SO Sequence 1358 BP; 408 A; 306 C; 295 G; 389 T; 0 U; 0 Other;

Query Match 100.0%; Score 1380; DB 8; Length 1398;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAAGCTGAACCATGTTAACTTATGACCTCGCTTAAATTTTAAATTATC 60  
 DB 1383 TAAAGCTGAACCATGTTAACTTATGACCTCGCTTAAATTTTAAATTATC 1324  
 QY 61 CATGCTAGCAAGCAGGTGCAAAACACAGCCCTGACGGCTTTGACGTGCGAA 120  
 DB 1323 CATGCTAGCAAGCAGGTGCAAAACACAGCCCTGACGGCTTTGACGTGCGAA 1264  
 QY 121 TACTGCCGATCAAGTGTACCACTGTATATTTAGTCCAAATGTGCATCACTTT 180  
 DB 1263 TACTGCCGATCAAGTGTACCACTGTATATTTAGTCCAAATGTGCATCACTTT 1204  
 QY 181 GAGTAAATCTTGCTCAATTAATCGGGGATCTGACGATTAATTAACACATATCCACCA 240  
 DB 1203 GAGTAAATCTTGCTCAATTAATCGGGGATCTGACGATTAATTAACACATATCCACCA 1144  
 QY 241 CTGATTAACATATGCTGACAGCTCACTAAATCTTGAAGTTCACCAAGCCCAAT 300  
 DB 1143 CTGATTAACATATGCTGACAGCTCACTAAATCTTGAAGTTCACCAAGCCCAAT 1084  
 QY 301 CAATGCCAGTGAATATGCTCAATTAACCGTGCACAGCCCATCAATGCTGATGTTGA 360  
 DB 1083 CAATGCCAGTGAATATGCTCAATTAACCGTGCACAGCCCATCAATGCTGATGTTGA 1024  
 QY 361 GCGGATTTTGATCTTTGAAATATGATTAATGCTGATCAATATCATCAATATA 420  
 DB 1023 GCGGATTTTGATCTTTGAAATATGATTAATGCTGATCAATATCATCAATATA 964  
 QY 421 TTGACAGCAGTGCAGTACCTTTGAAAGTCTTAAGGTAATCAAGATGCGATCATGGG 480  
 DB 963 TTGACAGCAGTGCAGTACCTTTGAAAGTCTTAAGGTAATCAAGATGCGATCATGGG 904  
 QY 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGTAAGTTAGTTGCC 540  
 DB 903 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGTAAGTTAGTTGCC 844

QY 541 TTGATTTTAAATGATGATGTTAAAGTTTGTGTTTGTATATATATGCAATACC 600  
 DB 843 TTGATTTTAAATGATGATGTTAAAGTTTGTGTTTGTATATATATGCAATACC 784  
 QY 601 ATCACCATCATCAATTAATTAATAATCTGCGTTTGTGAGTAACTAATTTGTGT 660  
 DB 783 ATCACCATCATCAATTAATTAATAATCTGCGTTTGTGAGTAACTAATTTGTGT 724  
 QY 661 GATTAATGCTGATTTTGGGCAACGCTTGTCAATGCTGATCAATGCTGCTAAAT 720  
 DB 723 GATTAATGCTGATTTTGGGCAACGCTTGTCAATGCTGATCAATGCTGCTAAAT 664  
 QY 721 GGCATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 DB 663 GGCATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604  
 QY 781 CAATTAATCAACCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB 603 CAATTAATCAACCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 544  
 QY 841 TGCCCTTGAAGTGCCTAAATTTGATATGCTCAAGCTGAAATCTGACAGCTCAAGC 900  
 DB 543 TGCCCTTGAAGTGCCTAAATTTGATATGCTCAAGCTGAAATCTGACAGCTCAAGC 484  
 QY 901 CACCAATATCAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 960  
 DB 483 CACCAATATCAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 424  
 QY 961 GACACCAACATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 DB 423 GACACCAACATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 364  
 QY 1021 TTGCGATTTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
 DB 363 TTGCGATTTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 304  
 QY 1081 TGCTTGAAGTATCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1140  
 DB 303 TGCTTGAAGTATCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 244  
 QY 1141 GCGGCGACAAATCTGCGGCTTCTTGGGTTAATGCGGCGCTGATGATGATGATG 1200  
 DB 243 GCGGCGACAAATCTGCGGCTTCTTGGGTTAATGCGGCGCTGATGATGATGATGATG 184  
 QY 1201 TGAGCTTAAGATGCTGCTGCAAGCTGCCAAATGAATGAATGCAATGCAATGCA 1260  
 DB 183 TGAGCTTAAGATGCTGCTGCAAGCTGCCAAATGAATGAATGCAATGCAATGCA 124  
 QY 1261 ATGCGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 123 ATGCGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 64  
 QY 1321 CAAAAAATTCACAGCTGACAGCTCCAGAGCTGCCAAACCGACGACGCAATTTTTTGA 1380  
 DB 63 CAAAAAATTCACAGCTGACAGCTCCAGAGCTGCCAAACCGACGACGCAATTTTTTGA 4

# RESULT 2

AAAF28548 standard; DNA; 96109 BP.

AAAF28548;

04-APR-2001 (first entry)

Genomic fragment #35.

Genomic library; bacteria; human upper airway; otitis media; sinusitis;

bronchopulmonary; endocarditis; meningitis; ss.

Moraxella catarrhalis.



PN MO200078968-A2.  
 XX 28-DEC-2000.  
 XX 16-JUN-2000; 2000MO-US016649.  
 XX 18-JUN-1999; 99US-0140121P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PI Lagace RE, Patterson C, Berg KL;  
 XX WPI; 2001-041427/05.  
 XX  
 PT Genomic library for identifying diagnostic and therapeutic compositions,  
 PT and for identifying virulence factors, regulatory elements and drug  
 PT targets, comprises Moraxella catarrhalis nucleic acids.  
 XX  
 PS Claim 1; Page 345-368; 545pp; English.  
 XX  
 CC The present invention relates to a Moraxella catarrhalis genomic library  
 CC comprising of a combination of 41 nucleic acid molecules (see AAF28514-  
 CC AAF28554). The library has a number of uses described in the  
 CC specification e.g. is useful for identifying diagnostic and therapeutic  
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large  
 CC aerobic, gram-negative diplococcus, normally found among the bacterial  
 CC flora of human upper airways. M. catarrhalis is known to cause acute,  
 CC localised infections such as otitis media, sinusitis and bronchopulmonary  
 CC infection and life-threatening, systemic diseases including endocarditis  
 CC and meningitis  
 XX  
 SO Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1380; DB 4; Length 96109;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAAGGTAAACCATTTAAACAACTTATCAGCTGCTCATTTATTTTAAATTGATC 60  
 DB 11357 TAAAGGTAAACCATTTAAACAACTTATCAGCTGCTCATTTATTTTAAATTGATC 11416  
 QY 61 CATGCTAGCAACAAGCAGTGCACAAAGCAAGCCCTAGCCGCTTGACTGCTGCCAA 120  
 DB 11417 CATGCTAGCAACAAGCAGTGCACAAAGCAAGCCCTAGCCGCTTGACTGCTGCCAA 11476  
 QY 121 TACTGCCCATCAAGTGTACCAACTGATATATTTAGATGCCAAATGTGCATCCTTT 180  
 DB 11477 TACTGCCCATCAAGTGTACCAACTGATATATTTAGATGCCAAATGTGCATCCTTT 11536  
 QY 181 GAGTAAATCTGTGCAATTAATCGGGGCACTTGAACGATTAATATACACTATCCACCA 240  
 DB 11537 GAGTAAATCTGTGCAATTAATCGGGGCACTTGAACGATTAATATACACTATCCACCA 11596  
 QY 241 CTGCATATACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCCAACCAAT 300  
 DB 11597 CTGCATATACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCCAACCAAT 11656  
 QY 301 CAATGCCAGTATGATATGCAATTAACCGTGCAGAGCCCAATCAATTGCTGCAATGTTGA 360  
 DB 11657 CAATGCCAGTATGATATGCAATTAACCGTGCAGAGCCCAATCAATTGCTGCAATGTTGA 11716  
 QY 361 GCGAGTGTGTATACCTTTGAAATCAATTAATATGCTTGTATCAATATCATCATATTA 420  
 DB 11717 GCGAGTGTGTATACCTTTGAAATCAATTAATATGCTTGTATCAATATCATCATATTA 11776  
 QY 421 TTACACAGCATGCGGTAGCCCTTTGAAAGCTTAAAGGTATCAAGATGGCATCCATGGG 480  
 DB 11777 TTACACAGCATGCGGTAGCCCTTTGAAAGCTTAAAGGTATCAAGATGGCATCCATGGG 11836  
 QY 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCATTAAGGCATTGAGTAGTTATGCTTGGC 540  
 DB 11837 AAGATTGGCAGCTGTGCTTAAGGCAAGTGCATTAAGGCATTGAGTAGTTATGCTTGGC 11896

QY 541 TTTGATTTTAAATGATGATTTGTTAAAGTTTGTATTTTATGATATATATCCATACC 600  
 DB 11897 TTTGATTTTAAATGATGATTTGTTAAAGTTTGTATTTTATGATATATATCCATACC 11956  
 QY 601 ATCACCATCATCAAAATTAATTAATTAATCTGCGTTGGTGGCTAAGTAAGCTATTTGTTGT 660  
 DB 11957 ATCACCATCATCAAAATTAATTAATTAATCTGCGTTGGTGGCTAAGTAAGCTATTTGTTGT 12016  
 QY 661 GATATGCTGTGATTTTGGGGCAACGCTTGTCAAGTGTGTCAAGCATTTGCTGGCTAAAT 720  
 DB 12017 GATATGCTGTGATTTTGGGGCAACGCTTGTCAAGTGTGTCAAGCATTTGCTGGCTAAAT 12076  
 QY 721 GGCATGCTTGTGTAGATGATGAGCGGTGTGGCAATTTTCAAAAATPAGCAATTTTGTGC 780  
 DB 12077 GGCATGCTTGTGTAGATGATGAGCGGTGTGGCAATTTTCAAAAATPAGCAATTTTGTGC 12136  
 QY 781 CAATTAATCCACATGCGATCGTCCGATCAAGATGATCAGCGAGAGATTTAAATTTGT 840  
 DB 12137 CAATTAATCCACATGCGATCGTCCGATCAAGATGATCAGCGAGAGATTTAAATTTGT 12196  
 QY 841 TGCCCTTGAAGTGCCTAAATTTGAGATATGCTCAAGCTAAATTTGACAGCTCAAGCAC 900  
 DB 12197 TGCCCTTGAAGTGCCTAAATTTGAGATATGCTCAAGCTAAATTTGACAGCTCAAGCAC 12256  
 QY 901 CACCAATATCATATTTTGGATGCTTAAATTAATCAAGTGCAGCGGTGCCAATATTAACGCC 960  
 DB 12257 CACCAATATCATATTTTGGATGCTTAAATTAATCAAGTGCAGCGGTGCCAATATTAACGCC 12316  
 QY 961 GACACCAACATATGCTGATGTTTTCATCTTGCCTACTAATGTCGTAAACAGTGTCT 1020  
 DB 12317 GACACCAACATATGCTGATGTTTTCATCTTGCCTACTAATGTCGTAAACAGTGTCT 12376  
 QY 1021 TTTGGATTTGAGCTGTGTATGCGCATGATGTGTGTGTTGCTTGAATTGATCTCG 1080  
 DB 12377 TTTGGATTTGAGCTGTGTATGCGCATGATGTGTGTGTTGCTTGAATTGATCTCG 12436  
 QY 1081 TGCTTGAATGATCAATGAACAATTGAACATGCGTATCAGCAGGAATGCTTGAAGCTTT 1140  
 DB 12437 TGCTTGAATGATCAATGAACAATTGAACATGCGTATCAGCAGGAATGCTTGAAGCTTT 12496  
 QY 1141 GCGGGGACAAATCTGCGGCTTCTTGGGTTAAATGCGGGGCTGATGATGATCTGATCGGC 1200  
 DB 12497 GCGGGGACAAATCTGCGGCTTCTTGGGTTAAATGCGGGGCTGATGATGATCTGATCGGC 12556  
 QY 1201 TGAAGCTTAAGATTCATCTGTCCAAAGTGCACAAATGAACAATGACGCTGTGGGAGTTT 1260  
 DB 12557 TGAAGCTTAAGATTCATCTGTCCAAAGTGCACAAATGAACAATGACGCTGTGGGAGTTT 12616  
 QY 1261 ATCTGCAAGATGAGATTTGGATTTACCATGATGATGCTTAACCTTTGTGACCGTATTTGAC 1320  
 DB 12617 ATCTGCAAGATGAGATTTGGATTTACCATGATGATGCTTAACCTTTGTGACCGTATTTGAC 12676  
 QY 1321 CAAAAAATTCACAGCTGCACGTCCAGAGCTGCCAAACGACGACGCAATATTTTGTGA 1380  
 DB 12677 CAAAAAATTCACAGCTGCACGTCCAGAGCTGCCAAACGACGACGCAATATTTTGTGA 12736

RESULT 3  
 ADL02987/c  
 ID ADL02987 standard; DNA; 1410 BP.  
 XX  
 AC ADL02987;  
 XX  
 XX 06-MAY-2004 (first entry)  
 XX  
 DE DNA encoding a M. catarrhalis protein #673.  
 XX  
 KM de; gene; Moraxella catarrhalis; infection.  
 XX  
 OS Moraxella catarrhalis.  
 XX  
 PN US6673910-B1.

```

PD 06-JAN-2004.
XX
XX 04-APR-2000; 2000US-00540236.
XX
XX 08-APR-1999; 99US-0128416P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL;
XX
XX MPI; 2004-178127/17.
XX
XX F-PSDB; ADL04907.
XX
XX
XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
XX PT preparing a composition for diagnosing, preventing or treating infection
XX PT caused by Moraxella catarrhalis.
XX
XX Disclosure; SEQ ID NO 673; 429bp; English.
XX
XX The invention relates to an isolated nucleic acid encoding an Moraxella
XX CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
XX CC composition for diagnosing, preventing or treating infection caused by
XX CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
XX CC catarrhalis protein.
XX
XX
XX Sequence 1410 BP; 409 A; 307 C; 302 G; 392 T; 0 U; 0 Other;
SQ
Query Match 99.3%; Score 1370.4; DB 12; Length 1410;
Base Local Similarity 99.6%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAAGCGTGAACCATGTTAAACAATTATCAGCTGCTCATTAATTTTAAATTGATC 60
Db 1392 TAAAGGTGACATGTTAAACAATTATCAGCTGCTCATTAATTTTAAATTGATC 1333
QY 61 CATGCTAGCAACAGAGTGACAAAAGCACAGCCCTAGCCGTGCTTGACTGCTGCCAA 120
Db 1332 CATGCTAGCAACAGAGTGACAAAAGCACAGCCCTAGCCGTGCTTGACTGCTGCCAA 1273
QY 121 TACTGCCGACGTAAGGTGACCAACCTGATATTTTATGATGCCAAATGTGCATCAGCTT 180
Db 1272 TACTGCCGACGTAAGGTGACCAACCTGATATTTTATGATGCCAAATGTGCATCAGCTT 1213
QY 181 GAGTAAATCTTGGCTCAATTTATCGGGGATCTTGACCGATTAATAATACACTATCCACCA 240
Db 1212 GAGTAAATCTTGGCTCAATTTATCGGGGATCTTGACCGATTAATAATACACTATCCACCA 1153
QY 241 CTGCATTAACATAGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT 300
Db 1152 CTGCATTAACATAGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAT 1093
QY 301 CAATGCAAGATGATATGTCATTAACCGTGCAGAGCCCATCAATTGCTGCATAGTTGA 360
Db 1092 CAATGCAAGATGATATGTCATTAACCGTGCAGAGCCCATCAATTGCTGCATAGTTGA 1033
QY 361 GCCGATGTTTGAACCTTTTGAATCATTTAAATAATGCTTGCTATCAATATCATATATA 420
Db 1032 GCCGATGTTTGAACCTTTTGAATCATTTAAATAATGCTTGCTATCAATATCATATATA 973
QY 421 TTACAGCGATGCGGTAGCCCTTTGAAGTCTTAAGGTATCAAGCATGGCATCCATGGG 480
Db 972 TTACAGCGATGCGGTAGCCCTTTGAAGTCTTAAGGTATCAAGCATGGCATCCATGGG 913
QY 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGAGATTAAGGCAATTGAGTATGCTTGGC 540
Db 912 AAGATTGGCAGCTGTGCTTAAGGCAAGTGAGATTAAGGCAATTGAGTATGCTTGGC 853
QY 541 TTGATTTTAAATGATGATGATGATTAAGTTTGTGTTTGAATGATATATGATCCATACC 600
Db 852 TTGATTTTAAATGATGATGATGATTAAGTTTGTGTTTGAATGATATATGATCCATACC 793
QY 601 ATGACCATCATCAATTAATAAATCTGCCGTTTGGTGGCTAAGTAACTATTTGTGT 660

```

```

Db 792 ATCACCATCATCAATAAATAAATAATCTGCCGTTTGGTGGCTAAGTAACTATTTGTGT 733
QY 661 GATTAATGATGATTTTGGGCAACGCTTGCTCATGTGATCAAGCATTTGCTGGCTAAATT 720
Db 732 GATTAATGATGATTTTGGGCAACGCTTGCTCATGTGATCAAGCATTTGCTGGCTAAATT 673
QY 721 GGCATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAAATAACGCAATTTTGTGC 780
Db 672 GGCATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAAATAACGCAATTTTGTGC 613
QY 781 CAATTAATCCACCATGCTGCTGCGCATCAAGATGATGACGGGAGAGATTTAAATTTGT 840
Db 612 CAATTAATCCACCATGCTGCTGCGCATCAAGATGATGACGGGAGAGATTTAAATTTGT 553
QY 841 TGCCCTTGAGGCTGCCAAATTTGAGATATGCTCAAGCTGAAATCTTACAGCTCAAGCAC 900
Db 552 TGCCCTTGAGGCTGCCAAATTTGAGATATGCTCAAGCTGAAATCTTACAGCTCAAGCAC 493
QY 901 CACCAATCCATATTTTGGATGCTTAATTAATTCAAAGTGAGGCGTGCCTAAATTAACGCGC 960
Db 492 CACCAATCCATATTTTGGATGCTTAATTAATTCAAAGTGAGGCGTGCCTAAATTAACGCGC 433
QY 961 GACACCAACAATCATGCTGCTGATGTTTGGCATTTGCTGCTACTAATATGCTTAACAGTGT 1020
Db 432 GACACCAACAATCATGCTGCTGATGTTTGGCATTTGCTGCTACTAATATGCTTAACAGTGT 373
QY 1021 TTTGGCATTTGAGGCTGCTGATGCGCATGATGGTGTGTTGCTTGTGATTAAGATCTCG 1080
Db 372 TTTGGCATTTGAGGCTGCTGATGCGCATGATGGTGTGTTGCTTGTGATTAAGATCTCG 313
QY 1081 TGGTTTGAAGTATCAATGAATGAATGAATGAAATGCTGATCAAGAGATGCTTGAAGCTTT 1140
Db 312 TGGTTTGAAGTATCAATGAATGAATGAATGAAATGCTGATCAAGAGATGCTTGAAGCTTT 253
QY 1141 GCGGGGACAATATCTGGGGCTTCTGGGTTAATGCCGGGGCTGATGATGATGATCGGCT 1200
Db 252 GCGGGGACAATATCTGGGGCTTCTGGGTTAATGCCGGGGCTGATGATGATGATCGGCT 193
QY 1201 TGAGCTTAAGAGTTCAGTCTGCTCAAGCTGSCAAATGACATGAAACGCTGTGGGAGTTT 1260
Db 192 TGAGCTTAAGAGTTCAGTCTGCTCAAGCTGSCAAATGACATGAAACGCTGTGGGAGTTT 133
QY 1261 ATCTGCCAAGTATGATGGATTTACATGATGATGCTTAATCTTGTGACCGTGAATGAC 1320
Db 132 ATCTGCCAAGTATGATGGATTTACATGATGATGCTTAATCTTGTGACCGTGAATGAC 73
QY 1321 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAACCGAGAGCGGCATATTTTGTGA 1380
Db 72 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAACCGAGAGCGGCATATTTTGTGA 13

```

RESULT 4  
ACA21070/c  
ID ACA21070 standard; DNA; 1344 BP.

ACA21070;  
19-JUN-2003 (first entry)

Prokaryotic essential gene #2727.

Antisense; ds; prokaryotic essential gene; cell proliferation;  
drug design; gene.

Acinetobacter baumannii.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002MO-US009107.

21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-0007283P.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW,  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR P-PSDB; ABU17200.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids, required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS  
 PS Claim 14; SEQ ID NO 8940; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation; (7) identifying a gene in an operon required for  
 CC the gene product or that has an activity against a biological pathway;  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data is one of the target  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1344 BP; 385 A; 254 C; 318 G; 387 T; 0 U; 0 Other;

Query Match 16.7%; Score 231; DB 8; Length 1344;  
 Best Local Similarity 52.4%; Pred. No. 1e-55;  
 Matches 721; Conservative 0; Mismatches 590; Indels 66; Gaps 7;

QY 2 AANGCGTGAACATGTTAACAACTTATCACTGCTCATTTATTTAAATGATCC 61  
 DB 1340 AAGCAATTTGACGACGACCAAACTGCTGACCAAGTCATTATTAACATTTAAACATTTCA 1281  
 QY 62 ATGCTTGACCAACAGGTGACAAAGACAGCCCTAGCGGTGTTTGAAGCGTGGCAAT 121  
 DB 1280 AATCTTGACATCTGGTGTAT-----AGCAATACACATCTTCAAGCTTGGTTTAC 1229  
 QY 122 ACTGCCGATCAAGTATCAACCTGATATTTTATAGTCCAAATGTGATCACTTTG 181  
 DB 1228 GTTGACACAGCTGACAGCTTTCTTAAGCGTTGTGATGATGTAATAATTGAGTGGCT- 1170  
 QY 182 AGTAAATCTTGCTCATATATGCGGCGATCTTGACCATTTAAATACACATATCAACCCAC 241  
 DB 1169 TGAATGCTTGTGCTGATGACCGGCGCATCTTCAACCAATCAATCACTTTGGCATAT 1110

QY 242 TGCATACATATGCTGACAGCTCACTAAATCTTGAACCTTTACCAAGCCCAACCAATC 301  
 DB 1109 TTTTCATATGAAAGATGTAAGAGAAATCTTGCCCTTTTGCCCTGTCACCTAAATA 1050  
 QY 302 AATGCCA-----GTGATATATGCAATAAACCGTGCACAGCCCATCAATTGCTCAATG 355  
 DB 1049 AGTGCAACCTTACCTTTTTCACCTTCATATGGCGACCTTAAGCCATCAATTTGCTGAAGT 990  
 QY 356 GTTGAGCCGATGTTTGTACTTTTGAATCATTTAAATATGCTTGTATCAATATCATCA 415  
 DB 989 GTAGCAACAACATTTGATCTTTAGATCATTTATTAACGACATCAACGACACATTTTA 930  
 QY 416 ATATATTCACAGCATGCGTAGCCCTTGAAGAGCTTAAGGATATCAAGCATGAGATCC 475  
 DB 929 ACGTACTCAAGAGATGCTCTATCTTTAAATCGTTTAAAGTTTCAAGCATTTGACTCC 870  
 QY 476 ATGGAAAGATGGCAGCTGTGCTTAAGCAAGTGCAGATTAAGCATTTGATGATATGC 535  
 DB 869 ATAGTAAACCAATTCCTCCCAATGCTAAACACAGCTAAAGCATTTGCTAGTGTGC 810  
 QY 536 TTGCTTTGATTTTAAATGATGATGATGTTGTTTATTAATGATATTAATGCC 595  
 DB 809 ATACCTTGAATATATAAATCCAGCTTTTAAATTAACCGCTTAACACAGACCAACCA 750  
 QY 596 ATACATCAACATCATCAATAATAATAATCTGCCGTTTGTGCTTAAGTAACTATTT 655  
 DB 749 AGGTACCGTGGCATCTCTTAATAAGCCATATTAAT-----TAAATCGTGGCT 698  
 QY 656 GTTGTATATATGTGATTTGGGGCAACGCTTGTCAGTGTGTCAAGCATTTGCTGCT 715  
 DB 697 TTAAACCAAGCTTTCATTTGATGTTGTGATCTGGAACCAAGTGGCGGCTTA----- 647  
 QY 716 AATTTGCAATCGCTTGTGATATGATGATGGGCTTTGGCAATTTTCAAAATPACCAATTT 775  
 DB 646 ----AAGCATGCTGTGGTAAATTAACCTTTTAAAGCCCTTGGAATAATGATGTTT 521  
 QY 776 TGTGCCAATTAATTCACATATGATGATGCTGCTGCAATGATGATGATGATTA 835  
 DB 590 GCTTGATGATATCCCAAGCATATTTCCATGACGCTTAATATGCTTCACTCATATTTGA 531  
 QY 836 ATTTGTCCTTGTAGCTGCAATTTGATGATGCTCAAGCTGAAACCTTGACAGCTCA 885  
 DB 530 AACACGCTAACCTCACCGTTTAAGTAAAGAGGTGTTCTTAACGTAAGCTTGAATCACTCA 471  
 QY 896 AGACCAACAATTCATATTTTGGATGCTTAATTAATCAAGGACGCGCAATATTA 955  
 DB 470 AGGACTATGATGCTC---TGGTGTATCTTTAAGTAATTTAATCCCTGTGACCAAGATTTG 414  
 QY 956 CCGCGACACCAACATATGCTGATGATGCTTGTGCAATCTGCTTCAATATGCTATCA 1015  
 DB 413 CCGCTTACGCAACTTTCTTACTGATCTTAAAGCATTTAAGCTGTTACA 354  
 QY 1016 GTGCTTTTGGCATTTGAGCCTGTGATGCGGATGATGTTGTGTTGCTTGAGTTAGA 1075  
 DB 353 GTACTTTTTCATTTGAACTGTATATGCAACATGCGCAATCACTGAGTGG----- 304  
 QY 1076 TCTGCTGCTTTAGTATCAATGAACATTTGCAATCTGCTATCAAGAAATGCTTGA 1135  
 DB 303 -----AGACGCAATTAATTTGATATCAACCCACACAGAAATAACCTTTA 261  
 QY 1136 GCTTTGGCGGCGCAATATCGGCTTCTTGGTTATGCTGGGGGCTGATGATGATCTGA 1195  
 DB 260 GCAATAGCTGCTGAATTTCCGCTATTTGTGTGCAAGAGCTTGCTTAAATAATCTCT 201  
 QY 1196 TCGGCTGAGCTTAAGGTTTCACTGTCACAGCTGCCAAATGACATGAACGCTGTGGGC 1255  
 DB 200 TCTGCTTGAATATTAATTTCTTGATCAACCTGCAAAACCTGGTTTAAACGCGACGCGA 141  
 QY 1256 AGTTATATGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315  
 DB 140 ATCTGATCTGCTCGGAGAGTGTGGGCGGAGTCCGTTTACACCAATCTTGATGCTTGT 81  
 QY 1316 TTGACCAAAAATATTCACAGCTGACAGCTGACAGCTGCCAAACCGACGAGCGATAT 1372

Db 80 TCATGACGAAATTACAGCAGAAACCTGATATTCCCAAGCCTGCTCAACTTTT 24

RESULT 5  
ADA29995/c  
ID ADA29995 standard; DNA; 1416 BP.  
XX  
XX ADA29995;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX DNA encoding Acinetobacter baumannii protein #1282.  
XX  
XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;  
XX vaccine; plant biocontrol agent.  
XX  
XX Acinetobacter baumannii.  
XX  
XX US6562958-B1.  
XX  
XX 13-MAY-2003.  
XX  
XX 04-JUN-1999; 99US-00328352.  
XX  
XX 09-JUN-1998; 98US-008701P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Breton G. Bush D;  
XX  
XX WPI; 2003-576092/54.  
XX  
XX P-PSDB; ADA4121.  
XX  
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
XX for diagnosing a bacterial disease, as components of antibacterial  
XX vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
XX plants.  
XX  
XX Example; SEQ ID NO 1282; 328bp; English.  
XX  
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.  
XX The A. baumannii nucleic acids and polypeptides are useful as reagents  
XX for diagnosing a bacterial disease, as components of antibacterial  
XX vaccines, as targets for antibacterial drugs, to detect the presence of  
XX A. baumannii and other Acinetobacter species in a sample, in screening  
XX or to inhibit A. baumannii infection, and as biocontrol agents for  
XX plants. The present sequence represents DNA encoding an A. baumannii  
XX protein.  
XX  
XX Sequence 1416 BP; 417 A; 262 C; 326 G; 411 T; 0 U; 0 Other;  
XX  
XX Query Match 16.4%; Score 226.2; DB 9; Length 1416;  
XX Best Local Similarity 52.1%; Pred. No. 2.5e-54;  
XX Matches 718; Conservative 0; Mismatches 593; Indels 66; Gaps 7;  
XX  
XX 2 AAGCGGAAACATGTTAACAACCTTACCTGCTCATTAATTTTAAATTGATCC 61  
XX 1409 AAGGAATTGACGACGACGACAACTGCTGACCAAGGATTAATTAACCTTTAAACATATCA 1350  
XX  
XX 62 ATGCTAGACACAGCAGGTGACAAAGACAGCCCTAGCGGTGCTTGACTGCGCCAAAT 121  
XX 1349 AAATCTGACATGCTGTGTAT-----AGCAATACACATCTTACGCTTGTTGTTTCA 1298  
XX  
XX 122 ACTGCCGATCAAGTGTACCAACCTGATATATTTTAATGATGCCAAATGTGATCACTTTG 181  
XX 1297 GTTGACACAGCTGACAGCTTTCTTTAAGCGTGTGATGTAAATTTTAAAGTTCGCT- 1239  
XX  
XX 182 AGTAATCTTGTCTCAATTATCGGGGACCTTGACACGATTAATTAACACATATCCACCAC 241  
XX 1238 TGAATGGCTTTGTTGATGACAGGCGCATCTTCAACCAATCAATACCACTTTGGCATAT 1179

QY 242 TGCATAATATGCTGACAGCTCATTAATCTTGACCTTTACCAAGCCCAACCAATC 301  
Db 1178 TTTTCATGTAAGAAAGTAAGGAAATACTGTGCTTTGCTCGTCCACTTAATAATA 1119  
QY 302 AATGCCA-----GTATGATATGCCATTAACCGGTGCCAAGCCCATCATTTGTGCAATG 355  
Db 1118 AGTGCAACCTTACCTTTTTCACCTTCATATGGCAGCACTTAAGCATTAATGTGTGCAAGT 1059  
QY 356 GTTAGCCGATGTTTGTACCTTTTGAATCATTTAAATATATGCTTGTCAATATCATCA 415  
Db 1058 GTAGCACCAACATTAATACCTTTAGAGTCATTTAATAAGGACATCATGACGGTTTA 999  
QY 416 ATATATCAACAGGATCGGTAGCCCTTTGAAGTCTTAAGGGTATCAAGCATGAGATCC 475  
Db 998 AGCTACTCACAGGGGTGCTAGTCTTAAATTTTAAAGTTTCAAGCATTTGACTCC 939  
QY 476 ATGGGAAGATGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGCATTTGATAGTTATGC 535  
Db 938 ATAGTAAACCAATTCCTTCCCAATGCTAAACAAAGCTAAAGCATTAATGCTATATGTGC 879  
QY 536 TTGCTTTGATTTTAAATGATGATTTGGTAAAGTTGTTTTTGTATGATTAATGCG 595  
Db 878 ATACCTTGAATATATAATCTGACTTTTAATTAACCGCTGTAAACGACGACCAACCA 819  
QY 596 ATACCATCAACCATCAATCAATAATAATAAATCGCCGTTTGTGCTTAAGTAACTATTT 655  
Db 818 AGCGTACCGTGGCATCTTTAAACGCCCATATGAT------TAATCCGGTGGCT 767  
QY 656 GTTGTATATAGTGTGATTTTGGGCAAGCTTTGTAGTGTGTCAAGCATTTGGCT 715  
Db 766 TTAAACCAAGCTTTGATTTGTGTGTCTGTAATGGAACAAAGTGCAGGCTTA----- 716  
QY 716 AAATGGCATGCTTGTGATGATGAGCGGTGTGCAATTTTCAAAATATACCAATTTT 775  
Db 715 ----AGCATCACTGGTTAATAATACACTTTTTCAGCCCTTGGAAATATGATGTTT 660  
QY 776 TGTGCCAATATATCAACCATGCGCATGTCCTCATGAATGATCAAGGAGATTTAA 835  
Db 659 GCTTGATGATATACCCAGCATATTTCCATGACGATTAATGATGCTTCACTCATATTAAGA 600  
QY 836 ATGTGTCCTTGTGACTGCCAAATTTGATATATGCTCAAGCTGAACCTGACACTCA 895  
Db 599 ACCATGCTACCTGACAGCTTTTAAGTAGAGGTGTTCTTAACGAAAGCTGATTAACCTCA 540  
QY 896 AGCACCACCAAAATCATTTTGTGATCGTTAATTAATTAATGACAGCGCTGCAATATTA 955  
Db 539 AGGACCAAGTACTGTG---TGATCTTTAAGTAATCTAAGCTGTGTGACCAAGGTTG 483  
QY 956 CGCGCAGACCAACATCATGCTGTGATGTTTGCATCTCGCTACTAATGTCTGTAACA 1015  
Db 482 CGGCTTACGCAACTTTCTTACCTGTCATCTTAAGCATTAAGCAATTAAGTGTTAACA 423  
QY 1016 GTGCTTTTGGCATTTAGCTGTGATGAGGAGATGATGTGTGTGTTGCTTGAAGTTAA 1075  
Db 422 GTACTTTTGGCATTTGAACTGTATATGTCACATGCGACATCACTGTC----- 373  
QY 1076 TCTGCTTTGAGTGTATCAATGAACAATTTGAACATGCTGATCAAGAAATGCTTGA 1135  
Db 372 -----AGGGGCAATTAATCTGATATCCGCCACACAAATTAACCTTTA 330  
QY 1136 GCTTTGGCGGCAACATTAATCTGCGCTTTTGGTTAATGCCGGGCTGATGATCTGA 1195  
Db 329 GCAATAGCTGCTGAAATTTCCGGTAATGTGTGACAAAGGCTGTGAAATAATATCTCT 270  
QY 1196 TGGGCTGAGCTTAAGTTTCACTGTGCCAAGCTGCCAAATGACAAATGAACATGACGCTGTGGCC 1255  
Db 269 TCTGCTTGAATATAATTTCTGATTAAGCTGACGCAAACTGTTTAAACCAAGCGGGA 210  
QY 1256 AGTTATCTGCCAAGGTGAGATTAACATCATGATGCTAATCTTTGTGACCGGTGA 1315  
Db 209 ATCTGATCGTGTGCGGGAAGTGTGGGGCGGGAATCGTTACAGCAACTTTGATGCTTGT 150  
QY 1316 TTGACCAAAAATTTACAGCTGACAGTCCAGAGCTGCCAAACCGACGAGCATAT 1372

Db 149 TCATGACGAAATTTACAGAGAAACACCTGATATTCCTCAAGCTCTCAACTTTT 93

RESULT 6  
AAS53453/c  
ID AAS53453 standard; DNA; 1314 BP.

XX AAS53453;

DT 13-FEB-2002 (first entry)

DE Haemophilus influenzae DNA for cellular proliferation protein #23.

XX Antisense, ds; prokaryotic cellular proliferation gene; antibiotic;

XX antibacterial; drug design.

XX Haemophilus influenzae.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU,

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU35594.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 7090; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the genes,

XX their use in the discovery of novel antibiotics, the essential genes,

XX themselves and the encoded proteins. The prokaryotes used are Escherichia

XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,

XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

XX useful for the identification of potential new targets for antibiotic

XX development. The antisense nucleic acids can also be used to identify

XX proteins used in proliferation, to express these proteins, and to obtain

XX antibodies capable of binding to the expressed proteins. The proteins can

XX be used to screen compounds in rational drug discovery programmes. The

XX antisense nucleic acid sequence is also useful to screen for homologous

XX nucleic acids which are required for cell proliferation in a wide variety

XX of organisms. The present sequence encodes an essential prokaryotic

XX cellular proliferation protein. Note: The sequence data for this patent

XX did not form part of the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

XX Query Match 11.6%; Score 160.4; DB 4; Length 1314;

XX Best Local Similarity 49.6%; Pred. No. 2.5e-35;

XX Matches 663; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

XX 1 TAAAGGTGAACCATTTAAACAATTCACCTGCTCATTTAAATTTTAAATTGATC 60

Db 1308 TAAACATTGAGCTAAATATGCGTAATTCCTCGCGGCTTTTCAAAAGAACAACTGATC 1249

Qy 61 CATGCTAGACAAAGAGGTGACAAAGACACACCTAGGCGTCTTGA CTGGTGCAC 120

Db 1248 GAGACTTGACAAAGACGGGACAAATATATGCA-----TATCTCGCTTGCAA 1201

Qy 121 TACTGCCGATCAAGTGTACCAACCTGATATTTTATAGATGCCAAATGTGATCACTTT 180

Db 1200 TGTGGGCGTAAATATTCATGCTTGTTCATTTGATGGAACA-----ATATCTTTG 1147

Qy 181 GAGTAATCTTGTCTCATTTATGCGGACATCTGACCGATTAATATACATATCCACCA 240

Db 1146 CGATGAATTTTGGACAGCGGACACATCTGACAAATATATATATATATGTGTG 1087

Qy 241 CTGCAATTAATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGGCCACCCAAAT 300

Db 1086 TTGATTAATTAATTCAGCTAAATTCGAAAATACAGCCCTTTTCGCTTCGCTTGCAA 1027

Qy 301 CAATGCCAGTGAATGATGCAATTAACCGTGCACAGCCCATCAATGTCTGCAATGTTGA 360

Db 1026 CAAATGCAATTTACCTCAATATA-----AGCCAGCCAAATGACGAACCTGTACT 976

Qy 361 GCGGATGTTTGTACCTTTTGAATCATTAATATATGCTTGTATCAATATATATATA 420

Db 975 CCCCAATTTGTGCTTTAGAGTCAATTAATCCAGCAATATAGCTTGTATGCACTTA 916

Qy 421 TTCAACAGCATGCGGATGCGCTTTGAAATCTTAAGGGATACAGCATGCAATCCATG 480

Db 915 TTGAAAACGATATATCTAACCTTTGAAATGACGAATGCGGTACGATTAATCTTAATT 856

Qy 481 AAGATTGCGAGCTGTGCTTAAAGCAAGTGCAGATTAAGCATTAAGGATTAATGCTTGC 540

Db 855 AATACCTATAGCTTGTGCGCAATGCTGTGCTGCAGAAATGTTCAATATATATGCGAGC 796

Qy 541 TTTGATTTTAAATGATGATTTGTAATGTTGTTTATGATATATATATGCAATACC 600

Db 795 AACCAATGATGCTTCTTCAAGTAAATCATCTTATCTTTACCAATTAATATGCTT 736

Qy 601 ATCAGCATATCAATTAATTAATTAATCTGC--CGTTGCTGCTATAGTATATTTGTTG 659

Db 735 GCCATTTTCAGTTTATGACCAATTAATCCGACTATTTTCCGAAAGAAACGATGTTT 676

Qy 660 TATATATGATGATTTGGGCAACGCTTGTCAATGTGTCAAGCATTTGCTGCTTAAT 719

Db 675 CGCTGATTTTCTTTTCCCAAAATGTCAGC----- 644

Qy 720 TGGCATGCTTGTGATGATGCGGTTTGGCAATTTTCAAAATATGCAATTTTGTG 779

Db 643 ---TATCTTCAATTTCAACACACCTTATGATATATGATTAATGCGTAATTTGCTT 587

Qy 780 CCAATTAATCCACATGCGCATGTCGCATCAAGATGATCAGCGAGAGATTTAAATG 839

Db 586 GGGGATTAATCTTAAATCAATATAGCGATATGATCTTCAAGTCAAGTTCMAACAG 527

Qy 840 TTGCCCTTGGAGCTGCCAATTTGATATGCTCAAGCTGAAACTTGCACGCTCAAGA 899

Db 526 TCGCTGCGGAGCTTTTAAGCTATATAGTTGTCTCAAGCTGAAAGCTGAAAGCTGAT 467

Qy 900 CCACCAATTCATATTTTGAATGTTAATTAATTAATGACAGCGCTGCCAATATTAACCC 959

Db 466 CATTAAGTTCACAATCTTCAATTC---AACATGACCAACCGGAAATCCCAATATTTCCCG 410

Qy 960 CGACACCAACATCATGCTGATGATGTTTGGCATCTGCGCTTCTATATGCTGAACGTGC 1019

Db 409 CCAATACCACTTTCAACACAGAGCTTGGCAATTTCAATTAACATAAGTATGATGCTGATC 350

Qy 1020 TTTTGGCATTTGAGCTGTGATGCGATGATGATGTTGTTGTTGCTTGAATGATCTC 1079

Db 349 TTTTACATTTGATCTGATATCCCAATTTGCTGTTGCTGCTGCGGCA----- 298

Qy 1080 GTGCTTGAATGATCAATGAACATGCAATGCTGATCAAGAGATGCTTGAAGCTT 1139

Db 297 -----GAATTAATCAATATGCGGATTAATCTTCACTCCGCTTTAA 257





```

Db      643 ---TATCTTATTTGTTCAACACACCTTATCTTACGATATATATTAATGCGTAATTTTGCTT 587
Qy      780 CCAATTAATTCACCATGCGCATCGTCCGATCAAGATATATCAAGGAGAGATTTAAATTG 839
Db      586 GGGGAAATATCTTTTAATTCATATAGGATTCATATATATCTTCAATGACCTTCAAGCAG 527
Qy      840 TTGCCCCCTTGAAGTGCCTCAATTTGATATATGCTCAAGCTGAAAAGTTCAGAGCTCAAGCA 899
Db      526 TCGCTGCGGAGCTTTTAACTATTAAGTTGTCTCAAGCTTAAAGTAAAGCTTAACTA 467
Qy      900 CCAACCAATTCATATTTTGGATGCTTAATTAATTCAGTGCAGGCGTCCAAATTAACCG 959
Db      466 CATTAAGTTTCAATCTTCAATTC---AACATGACAAAGGGGGAATCCCAATATTTCCG 410
Qy      960 CGACACCAACATCAATGCTGCTGATGTTTGGCTTCCATCTGCTACTAATGTGTAACTGTC 1019
Db      409 CCAATACCAATTCATCAACCAAGAGCTTCCGCAATTCATTAAGCTAAAGTATACGGTAC 350
Qy      1020 TTTTGGATTTGAGCGCTGATGCGGATGATGCTGTTGTTGCTTGAAGTATCTC 1079
Db      349 TTTTACCATTTGAACCTGTAATCCCAATGCGCTTTGCTGCGCGCA----- 238
Qy      1080 GTGCTTTGAGTGTATCATATGACATTTGAATGACATGCTGATCAGAGATGCTTGAAGCTT 1139
Db      297 -----GAATATTCATATTCGCGATTAATCTTCACTCCGCTTAA 257
Qy      1140 TGGCGGCGACATATCTGCGGCTTCTTGGTTAATCCGCGGCTGATGATGATCTGATCGG 1199
Db      256 GTGCGGTTTGAATTTCTGCTGTTTAAACCGCAAGCCCTGCGGCTTAATTAACATCAATCGC 197
Qy      1200 CTGAGCTTAAGATTCATCTGCTCAAGCTGCGAAATGACAAATGACATGACCCCTGTGGCAGTT 1259
Db      196 TTTCAGATTAACCATCTCTGATTTTAACTACCAAGTATGAAGAGGATATTTTGAAGAGATT 137
Qy      1260 TATCTGCAAGATGATGATTTGGGATTTACATGATGATGATGATGATGATGATGATGATGAT 1319
Db      136 TATCAATATACAGATGATGATTTTTCGATATCAATCAACAGAAATTAAGCTGTGGGATA 77
Qy      1320 CCAAAAATTCACAGCTGACATGCTGACAGCTGCCCAACCGACGACGCAATATTTT 1377
Db      76 AGAGATTAATTCACACAAAGAACCTGTGTTGGCAAGCCGATGATGATGATGATGATGATGAT 19

RESULT 8
ADL46439/c
ID      ADL46439 standard; DNA; 1314 BP.
XX
AC      ADL46439;
XX
DT      20-MAY-2004 (first entry)
XX
DE      aspartate semialdehyde dehydrogenase DNA #1.
XX
XX
XX      ds; gene; antibacterial;
XX      UDP-N-acetylglucosamine 1-carboxyvinyl transferase-1;
XX      CTP: CMP-3-deoxy-D-manno-ocutulosonate transferase;
XX      UDP-N-acetylmuramylalanine-D-glutamate-2-6-diaminopimelate ligase;
XX      D-alanine-D-alanine adding enzyme; D-alanine-D-alanine ligase;
XX      UDP-N-acetylpyruvoylglycosamine reductase;
XX      UDP-N-acetylglucosamine pyrophosphorylase;
XX      UDP-N-acetylmuramylalanine-D-glutamate ligase;
XX      DP-N-acetylmuramylalanine-D-glutamate semialdehyde dehydrogenase;
XX      UDP-N-acetylmuramylalanine-D-glutamate; X-ray diffraction analysis.
XX
XX      Haemophilus influenzae.
XX
XX      W02003087353-A2.
XX
XX      23-OCT-2003.
XX
XX      08-APR-2003; 2003MO-CA000481.

```

```

PR      08-APR-2002; 2002US-0370899P.
PR      08-APR-2002; 2002US-0370915P.
PR      09-APR-2002; 2002US-0371107P.
PR      09-APR-2002; 2002US-0371185P.
PR      31-MAY-2002; 2002US-0385426P.
PR      06-JUN-2002; 2002US-0386283P.
PR      01-AUG-2002; 2002US-0400348P.
PR      06-NOV-2002; 2002US-0424395P.
PR      08-NOV-2002; 2002US-0425200P.
PR      24-DEC-2002; 2002US-0436345P.
PR      24-DEC-2002; 2002US-0436349P.
PR      26-DEC-2002; 2002US-0436568P.
PR      27-DEC-2002; 2002US-0436675P.
PR      27-DEC-2002; 2002US-0436734P.
PR      27-DEC-2002; 2002US-0436885P.
PR      27-DEC-2002; 2002US-0436889P.
PR      27-DEC-2002; 2002US-0436893P.
PR      27-DEC-2002; 2002US-0436900P.
PR      30-DEC-2002; 2002US-0437013P.
XX
XX      (AFPI-) AFFINITUM PHARM INC.
XX
XX      Edwards A, Dharamsi A, Vedadi M, Domagala M, Houston S, Awrey D;
XX      Beattie B, Mansoury K, Ouyang H, Vallee F, Richards D, Neherly K;
XX      Virag C, Buzadzija K, Pinder B, Alam MZ, Tai M, Canadian V;
XX      Kanagarajah D, Thalakada R;
XX
XX      WPI; 2003-865361/80.
XX      P-PSDB; ADL46440.
XX
XX      New recombinant bacterial enzymes involved in cell membrane biogenesis,
XX      useful for designing potential antibacterial agents.
XX
XX      Claim 467; SEQ ID NO 157; 407bp; English.
XX
XX      The invention relates to isolated, recombinant polypeptides (I) that have
XX      at least one activity of specified bacterial enzymes involved in cell
XX      membrane biogenesis. (I) are: UDP-N-acetylglucosamine 1-carboxyvinyl
XX      transferase-1 of Streptococcus pneumoniae (S.p), Pseudomonas aeruginosa
XX      (P.a.) or Staphylococcus aureus (S.a.); CTP: CMP-3-deoxy-D-manno-
XX      octulosonate transferase of Escherichia coli (E.c.) or Haemophilus
XX      influenzae (H.i.); UDP-N-acetylmuramylalanine-D-glutamate-2,6-
XX      diaminopimelate ligase of P.a.; D-alanine-D-alanine adding enzyme of S.a.
XX      or P.a.; D-alanine-D-alanine ligase of Enterococcus faecalis (E.f.); UDP-N-
XX      acetylpyruvoylglycosamine reductase of P.a. or H.i.; UDP-N-
XX      acetylmuramylalanine-D-glutamate ligase of E.f., H.i. or S.a.; UDP-N-
XX      acetylmuramylalanine-D-glutamate ligase of E.f. or H.i.; DP-N-
XX      acetylmuramylalanine-D-glutamate-2-6-diaminopimelate ligase of E.c.;
XX      dehydrogenase of H.i. and UDP-N-acetylmuramylalanine-D-glutamate (sic) of
XX      H.i. Crystalline (I) are used to determine (by X-ray diffraction
XX      analysis) the structural coordinates of (I), and these then used to
XX      design modulators of (I), potential therapeutic agents for treating
XX      diseases caused by the specified bacteria. This sequence represents a DNA
XX      of the invention.
XX
XX      SO      Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;
XX
XX      Query Match      11.6%; Score 160.4; DB 11; Length 1314;
XX      Best Local Similarity 49.6%; Pred. No. 2,5e-35;
XX      Matches 663; Conservative 0; Mismatches 606; Indels 89; Gaps 7;
XX
Qy      1 TAAAGGTGATACCATTTTAACTATCACTGCTCATATATATTTTAAATGATC 60
Db      1308 TAAACATTTAGCTTAATGCGTAAATTTCTTGGCGGCTTTCAAAAGAAACGATC 1249
Qy      61 CATGCTAGCAAGACAGTGTGACAAAGACAGCGCTTGAATGCTGCTGCCAA 120
Db      1248 GAGACTTGTGACAGACGAGGAGCAATATATCA-----TATCTCCGCTTGCAA 1201
Qy      121 TACTGCCGATCAAGTGTACCAACTGATATATTTAGATGCCAAATGTGATCATT 180
Db      1200 TGTGGGCGTAAAAATTTATGCTGTGTTCCATTTGATGACAA-----ATAACTTG 1147

```



```

OY 181 GAGTAAATCTTGCTCAATTATCGGGGCACTGTGACCGATTAATACACTATCCACCA 240
Db 1146 CGATGAAAATTTTTCAGACGCGGACCATCTCGACCAAAACATTAACAAATTAATGTGG 1087
OY 241 CTGCAATACATAGCTGACAGCTCAGTAAATCTTGAAGCTTTACCAAGCCCAACCAAT 300
Db 1086 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1027
OY 301 CAATGCGAGTATGATATGCGATTAACCGTCCAGGCCATCAATGCTGCAATGGTTGA 360
Db 1026 CAATGCAATTAATCCCAATATAT-----AAGCCAGCCCAATGACCAACTGACT 976
OY 361 GCGGATGTTTGAACCTTTGAATCATTAAATATGCTTGTATCAATATCAATATATA 420
Db 975 CCCACATTTGTTGCTTTAGATCATTAATCAACGAATGCCATTAAGCTTGATGACTAA 916
OY 421 TTACAGCGATGGGATGAGCCCTTTGAAAGTCTTAAGGATCAAGCATGAGTATGCTTGC 480
Db 915 TTGAAAAAGATGATCTAACCCTTTGAATGACGAAGCGGTACGAATTAATCTAAT 856
OY 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTGAATGAAGCATGAGTATGCTTGC 540
Db 855 AATACCTATAGCTTGGCCATGCTGTGCTGCCAAAATGTCATATTAATTAAGGAC 796
OY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 795 AACCAATGATGCTTCTTCAAGATTAATCACTTCACTTCACTTCACTTCACTTCACT 726
OY 601 ATGACATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 659
Db 735 GCCATTTTCACTTTTGAACCAATTAATCCGACTAATTTTCCGCAAAAGAAACGATG 676
OY 660 TGATTAATGATGATTTGGGGCAAGCTTGTCAAGTGTGTCAAGCATCTGCTGCTAAT 719
Db 675 CGTTGATTTTGTGTTTCCCAAAAGTCAAGC----- 644
OY 720 TGGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
Db 643 ---TATCTTATGTTTCAACACACTTACTTATGATATGATTAATGCGTAATTTGCT 587
OY 780 CCAATTAATCCACATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 839
Db 586 GGGCATTAATCTTAAATCAATAGGATCAATGATCTTCAATGATCACTTAAGCAG 527
OY 840 TTGCCCCCTGAGCTGCAATTTGAGATATGCTCAAGCTGAACCTTGACAGCTCAAGCA 899
Db 526 TCGCTGCGGAGCTTTTAACCTTAATGATGCTCAAGCTGAACCTTAAGCTTAAGTA 467
OY 900 CCAACCAATCAATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 959
Db 466 CATTAAGTTCACATCTTCAATTC---AACATGACAAAGCGGAAATCCCAATTTCC 410
OY 960 CGACACCAACATCAATGCGCATGTTTGGCATCTGCGCTACTAATGCTGAAGTGC 1019
Db 409 CCAATCAACATTTCAACGACGAGCTTGGCAATTCATTAACCTTAAGTGAAGTAC 350
OY 1020 TTTTGGCATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
Db 349 TTTTACCATTTGAACTGATATCCCAATGCTGCTTGTGCTGCGCGGCA----- 288
OY 1080 GTGCTTGAAGTATCAATGAACATTTGAACATGCTGATCAAGAACTCTTGAGCTT 1139
Db 297 -----GAATTAATTCATTCGCGATTAATCTTCACTCCGCTTTAA 257
OY 1140 TGGCGGGAACAATCTGCGGCTTGTGATTAATGCGGGGCTGATGATGATGATGATG 1199
Db 256 GTGCGGTTGAATTTCTGCTGTTTTCACCGCAAGCCTTGAGCTTAATCAATATCTGC 137
OY 1200 CTGAGCTTGAAGTCACTGTCACAGCTGCAAAATGACAAATGACAGCTGTGGGAGTT 1259
Db 196 TTTCAAGTAAACATTCCTGATTTAACTACACAGATGAAGAGGATTTTGAAGAAATT 137
OY 1260 TATCTGCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319

```

```

Db 136 TATCAATACAGATGATTTTTCAGTATCAATCAACGAATATTAGCTGTGGATA 77
OY 1320 CCAAAAATTTCAAGCTGACAGTCCAGAGCTGCCCAACGACGACGATATTTT 1377
Db 76 AGGATTAATCCACACAAAGACCTGTTTGGCAAGCCGAGATGATTAATTTT 19

RESULT 9
AEC10868/C
ID AEC10868 standard; DNA; 1314 BP.
XX
AC AEC10868;
XX
DT 20-OCT-2005 (first entry)
XX
DE Haemophilus influenzae MURD DNA.
XX
KM protein purification; antibacterial; antimicrobial; infection;
KM drug screening; gene; ds; UDP-N-acetylmutramoylalanine-D-glutamate ligase.
XX
OS Haemophilus influenzae.
XX
PN US2005181388-A1.
XX
PD 18-AUG-2005.
XX
PF 04-OCT-2004; 2004US-00958216.
XX
PR 02-APR-2002; 2002US-0369511P.
PR 04-APR-2002; 2002US-0369617P.
PR 04-APR-2002; 2002US-0370102P.
PR 08-APR-2002; 2002US-0370778P.
PR 08-APR-2002; 2002US-0370792P.
PR 08-APR-2002; 2002US-0370820P.
PR 08-APR-2002; 2002US-0370859P.
PR 08-APR-2002; 2002US-0370899P.
PR 08-APR-2002; 2002US-0370915P.
PR 09-APR-2002; 2002US-0371067P.
PR 09-APR-2002; 2002US-0371107P.
PR 09-APR-2002; 2002US-0371140P.
PR 09-APR-2002; 2002US-0371185P.
PR 31-MAY-2002; 2002US-0385089P.
PR 31-MAY-2002; 2002US-0385426P.
PR 04-JUN-2002; 2002US-0385751P.
PR 05-JUN-2002; 2002US-0386018P.
PR 05-JUN-2002; 2002US-0386367P.
PR 05-JUN-2002; 2002US-0386548P.
PR 05-JUN-2002; 2002US-0386553P.
PR 05-JUN-2002; 2002US-0386566P.
PR 05-JUN-2002; 2002US-0386577P.
PR 06-JUN-2002; 2002US-0386283P.
PR 06-JUN-2002; 2002US-0386390P.
PR 06-JUN-2002; 2002US-0386430P.
PR 06-JUN-2002; 2002US-0386601P.
PR 06-JUN-2002; 2002US-0386826P.
PR 06-JUN-2002; 2002US-0386869P.
PR 31-JUL-2002; 2002US-0399972P.
PR 01-AUG-2002; 2002US-0400148P.
PR 05-NOV-2002; 2002US-0424053P.
PR 06-NOV-2002; 2002US-0424380P.
PR 06-NOV-2002; 2002US-0424395P.
PR 08-NOV-2002; 2002US-0425086P.
PR 08-NOV-2002; 2002US-0425200P.
PR 24-DEC-2002; 2002US-0436243P.
PR 24-DEC-2002; 2002US-0436288P.
PR 24-DEC-2002; 2002US-0436345P.
PR 24-DEC-2002; 2002US-0436349P.
PR 26-DEC-2002; 2002US-0436566P.
PR 26-DEC-2002; 2002US-0436567P.
PR 26-DEC-2002; 2002US-0436568P.
PR 27-DEC-2002; 2002US-0436575P.
PR 27-DEC-2002; 2002US-0436708P.

```



|            |   |   |      |
|------------|---|---|------|
| Oy         | 1260  | TATCTCCGAAGTATGATTTGGGATATACCATGCATATCTAACTTTGGACCGGATTTGA    | 1337 |
|            |   |   |      |
| Db         | 136   | TATCAATATCCAGTAGATTTTTCGAGTATCAATATACACGAAATATTAAGCTTTGGGATTA | 77   |
|            |   |   |      |
| Oy         | 1320  | CCAAAAATTACACAGCTGACAGTCCAGAGCTGCCAAACGACGACGCGCATATATTTTT    | 1377 |
|            |   |   |      |
| Db         | 76  | AGAGATATATTCACACAGAAAGAAAGACTGTTTGGCCAAAGCCGATGATATATATTTTT   | 19   |
|            |   |   |      |
| RESULT 10  |   |   |      |
| AE331251/c |   |   |      |
| ID         | AE331251  | standard; DNA; 2925 BP.                                       |      |
| XX         |   |   |      |
| AC         | AE331251/   |   |      |
| XX         |   |   |      |
| DT         | 09-FEB-2006   | (first entry)   |      |
| XX         |   |   |      |
| DE         | Haemophilus influenzae  | murD coding sequence.   |      |
| XX         |   |   |      |
| KW         | vaccine; antibacterial; murD; coding sequence; ds.                        |   |      |
| XX         |   |   |      |
| OS         | Haemophilus influenzae.   |   |      |
| XX         |   |   |      |
| PN         | US2005272089-A1.  |   |      |
| XX         |   |   |      |
| PD         | 08-DEC-2005.  |   |      |
| XX         |   |   |      |
| PF         | 01-AUG-2005; 2005US-00194246.   |   |      |
| XX         |   |   |      |
| PR         | 19-OCT-2001; 2001US-0345438P.   |   |      |
| XX         |   |   |      |
| PR         | 21-OCT-2002; 2002US-00274586.   |   |      |
| XX         |   |   |      |
| PA         | (MOT/) MOT J E.   |   |      |
| XX         |   |   |      |
| PA         | (TREP/) TREP D C M.   |   |      |
| XX         |   |   |      |
| PA         | (ARV/) ARVIDSON S.  |   |      |
| XX         |   |   |      |
| P1         | Mott JE, Trepod CM, Arvidson S;   |   |      |
| XX         |   |   |      |
| XX         | WPI; 2006-037204/04.  |   |      |
| DR         | P-PSDB; AE331355.   |   |      |
| XX         |   |   |      |
| PT         | Novel isolated polypeptide critical for survival of Haemophilus species,  |   |      |
| XX         |   |   |      |
| PT         | useful for reducing virulence of H. influenzae to be used as vaccine      |   |      |
| XX         |   |   |      |
| PT         | composition.  |   |      |
| XX         |   |   |      |
| PS         | Example 3; SEQ ID NO 220; 158bp; English.                                 |   |      |
| XX         |   |   |      |
| CC         | The invention relates to an isolated polypeptide (I) critical for the     |   |      |
| CC         | survival of Haemophilus sp. comprising an amino acid sequence having at   |   |      |
| CC         | least 95% structural similarity with an amino acid sequence chosen from   |   |      |
| CC         | AE331317, AE331330, AE331332, AE331335, AE331345, AE331353, AE331354 and  |   |      |
| CC         | AE331358. Also described: (1) an isolated polynucleotide (II) critical    |   |      |
| CC         | for the survival of Haemophilus sp. comprising a nucleotide sequence      |   |      |
| CC         | having at least 95% structural similarity with nucleotide sequence chosen |   |      |
| CC         | from a coding sequence in AE331062, AE331130, AE331140, AE331155,         |   |      |
| CC         | AE331211, AE331241, AE331263 and AE331231, and its complements, and       |   |      |
| CC         | encoding (I); (2) identifying (MI) an agent that binds a polypeptide,     |   |      |
| CC         | involves combining a polypeptide and an agent to form a mixture, where    |   |      |
| CC         | the polypeptide is encoded by a coding sequence comprising a nucleotide   |   |      |
| CC         | sequence having at least 95% structural similarity with a nucleotide      |   |      |
| CC         | sequence chosen from AE331062, AE331130, AE331140, AE331155, AE331211,    |   |      |
| CC         | AE331241, AE331246, AE331263 and AE331231, or has at least 95% structural |   |      |
| CC         | similarity with an amino acid sequence chosen from AE331317, AE331330,    |   |      |
| CC         | AE331332, AE331335, AE331353, AE331354 and AE331358, and                  |   |      |
| CC         | determining whether the agent binds the polypeptide; (3) decreasing the   |   |      |
| CC         | growth rate of a microbe, involves combining a microbe with an agent that |   |      |
| CC         | binds to a polypeptide encoded by a coding sequence comprising a          |   |      |
| CC         | nucleotide sequence chosen from AE331062, AE331130, AE331140, AE331155,   |   |      |
| CC         | AE331211, AE331241, AE331246, AE331263 and AE331231; (4) making (M2) an   |   |      |
| CC         | H. influenzae with reduced virulence, which involves altering a coding    |   |      |
| CC         | sequence in an H. influenzae to comprise a mutation, the non-mutagenized  |   |      |
| CC         | sequence in a nucleotide sequence chosen from AE331062,                   |   |      |

| CC                        | AE831130, AE831140, AE831155, AE831211, AE831241, AE831246, AE831263 and  |
|---------------------------|---|
| CC                        | AE831231, and determining if the H. influenzae comprising the mutation    |
| CC                        | has reduced virulence compared to an H. influenzae that does not comprise |
| CC                        | the mutation; (5) an H. influenzae obtained by (M2); and (6) a vaccine    |
| CC                        | composition comprising the H. influenzae. (1) is useful for reducing the  |
| CC                        | virulence of H. influenzae to be used as a vaccine composition, and for   |
| CC                        | decreasing the growth rate of a microorganism. The present sequence       |
| CC                        | represents a H. influenzae mud coding sequence, which is used in the      |
| CC                        | exemplification of the present invention.                                 |
| XX                        | Sequence 2925 BP; 796 A; 454 C; 640 G; 1035 T; 0 U; 0 Other;              |
| XX                        |   |
| Query Match               | 11.6%; Score 160.4; DB 15; Length 2925;                                   |
| Best Local Similarity     | 49.6%; Pred. No. 3,76-35;   |
| Matches 683; Conservative | 0; Mismatches 606; Indels 89; Gaps 7;                                     |
| QY                        | 1 TAAAGCGTGAACCATGTTAAACAACTTATCATCCCTGCTCATATATTTTAAATTATC 60            |
| DB                        | 2240 TAAACATTGAGCTAAATGCGTAAATTTCTTGCCCGCGCTTTTCAAMAMAGCAAACTGATC 2181    |
| QY                        | 61 CATGCTGACACAGCAGGTGACAAAGACACAGCCCTTAGCGGTGCTTGACTGGCTGCCAA 120        |
| DB                        | 2180 GAGACTTGACACAGCAGCGCACATTAATACCA-----TATCTCGCTTGCCAA 2133            |
| QY                        | 121 TACTGCCGATCAAGTGAACCAACCTGATATATTTTAGATGCCAATGTGCATCACCTT 180         |
| DB                        | 2132 TGTGGGGGTAAATTTCTATCGCTTGTTCATGTATCGAACAA-----ATATCTTGT 2079         |
| QY                        | 181 GAGTAAATCTTGCTCATATATCGGGGATCTTGACCATTAATAATACACTATTCACACCA 240       |
| DB                        | 2078 CGATGAATAATTTGCCAAGCAGCGACCACTTCGACCAAAACAATPACAAATATGTGTGG 2019     |
| QY                        | 241 CTGCATACATATGCTGACAGCTACCTPAAATCTTGACCTTTACCAAGCCACCACCAAA 300        |
| DB                        | 2018 TTGATTTATTAATTCAGCTATATCTGAATAATACGCCCTTTTTCGTCGCGCTACCA 1959        |
| QY                        | 301 CAATGCCAGTATATATGCCATTAACCCGTGCCAAGCCCATCAATTGCTGCATGTGTGA 360        |
| DB                        | 1958 CAAATGCAATTTACCTCAATATA-----AAGCCACGCAATGCACCAACTGTACT 1908          |
| QY                        | 361 GCCGATGTTGTACCTTTTGTGAATATTTAAATATGCTTGTATCATATATCATCAATATA 420       |
| DB                        | 1907 CCCCAATTTTGTCTTTAGATCATATTAATCCAAAGCAATGCAATGCAATGATGATGACATA 1848   |
| QY                        | 421 TTCAAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGTATCAAGCATGCAATGCATCATGG 480      |
| DB                        | 1847 TTGAAAAAGATGATCTAACCTTTGAAATATGCGAAGCGGTGCGAATTTGAATCTTAATT 1788     |
| QY                        | 481 AAGATGTGCGAGCTGTGCTTAAGCAAGTGCAGATPAAAGCAATGTGATGTATATGCTTGC 540      |
| DB                        | 1787 AATACCTTATGCTGTGTCAGATGCTGTGTGCTGCCAAATGTTCAATATTAATTAAGGCACC 1728   |
| QY                        | 541 TTGTATTTTAAATGATGATGTGTGTAAGTTGTTGTTTTTGTGATGATATAATGCAATACC 600      |
| DB                        | 1727 AACCAATGTACTCTTTCACAAAGTAAATATCACTTATCTTTTACATTAATATATGCTT 1668      |
| QY                        | 601 ATGACCATCATCAATTAATTAATAATCTGC-CGTTTGTGTGCTTAAGTAACTATTTGTG 659       |
| DB                        | 1667 GCCATTTTCACTTTTGGCCAAATATATCCGACTATTTTCCGCAAAAAGAAACGGATATTTT 1608   |
| QY                        | 660 TGATAATGTGTGATTTGGGGCAACGCTTGTCAGTGTGTGTCAAGCATTTGCTTGCTAAAT 719      |
| DB                        | 1607 CGCTGATTTTCTTTTCCCAAAAGTCAGC----- 1576                               |
| QY                        | 720 TGGCATGCTTTGGTGAATGATGTGGCGGCTTTGGCAATTTTCAAAAATACGCAATTTTGTG 779     |
| DB                        | 1575 ---TATCTTCAATGTTCAACACACTTACTTTAGCATATATATAATCCGTAATTTTGCTT 1519     |
| QY                        | 780 CCAATATATCACAATGCAATCGGCGGATCAAGATATATCAGCGGAGATTTAAATTTG 839         |
| DB                        | 1518 GCGGATATCTTTTAAATCATATAGCAATCATATATATCTTTCAGTACGTTCAAGACAG 1459      |
| QY                        | 840 TTGCCCCCTTGAAGTCAATTTGAGATATGCTTACAGTAAACCTTGACAGCTCAAGCA 899         |

```
Db      1458 TCGCTGCGGAGCTTTTAAGCTAATAGTGTCTCAAGCTGAAAAGCTAGTA 1399
Qy      900 CCACCAATTCATATTTTGGATGTTAATATTCAGTCAAGCGCTGCCAATATACCG 959
Db      1398 CATTAAGTTCACATCTTCAATTC---AACATAGCAAAAGCGGAATCCCAATATTTCCGC 1342
Qy      960 CGACACCAACATCATGCGCTGCATGTTTGGCATCTCGCTACTGAATGCTAACGTCG 1019
Db      1341 CCAATACCAATTTTCAACGACGAGCTTTCGCAATTCATTAACATTAAGTATGCGTAC 1282
Qy      1020 TTTTGGCATTTGAGCTGTGATGCGATGATGTTGTGTGTTTGTCTTGAATGATCTC 1079
Db      1281 TTTTACCATTTGAACTGTAAATCCCAATGCGCTTTGTGCGTGGCGGCA----- 1230
Qy      1080 GTGCTTTGAGTGTATCAATGAACAATGAACATGCTGATCAACGAAATGCTTGAAGCTT 1139
Db      1229 -----GAAATTAATTCAAATATCGCGATTTACTTCCACTCCGCTTTAA 1189
Qy      1140 TGGCGGCGACAACTACTGCGGCTTCTGGGTTAATGCCGGGCTGATGATGATCGATCGG 1199
Db      1188 GTGCGGTTTGAATTTCTGTTGTTTACGCGAAGCCTGCGCTAATTAACAATATTCGC 1129
Qy      1200 CTGAGCTTAAAGATTCATGCTGCCAAGCTGCCAAATGACATGAACGCTGTGGGAGTT 1259
Db      1128 TTTCAAGTAACTCTGTATTTAACTACAGATGAAGAGGATTTTGAAGAAATT 1069
Qy      1260 TATCTGCCAAGTGAAGTGGGATTAACATGATGATGATGATGATGATGATGATGATGAT 1319
Db      1068 TATCAATATCACGATGAGATTTTTCGAGTATCAATCAACAGAAATTAAGCTGTGGGATA 1009
Qy      1320 CCAAAAATTCACAGCTGACGATGCCAGCTGCCAAACGAGAGCGGCAATTTTTT 1377
Db      1008 AGAGTATATCAACAAGAAAGACGTGTTTGGCAAGCCGCGATGATGATATTTTT 951
```

## RESULT 11

AAAT42063\_11/c  
Continuation (12 of 19) of AAAT42063 from base 1100001 (Haemophilus influenzae comp)ete g  
WP Sequence Split Info 19 fragments LOCUS AAAT42063 Accession Aat42063

```
WP Fragment Name      Begin      End
WP AAAT42063_00      1      110000
WP AAAT42063_01     100001   210000
WP AAAT42063_02     200001   310000
WP AAAT42063_03     300001   410000
WP AAAT42063_04     400001   510000
WP AAAT42063_05     500001   610000
WP AAAT42063_06     600001   710000
WP AAAT42063_07     700001   810000
WP AAAT42063_08     800001   910000
WP AAAT42063_09     900001  1010000
WP AAAT42063_10    1000001  1110000
WP AAAT42063_11    1100001  1210000
WP AAAT42063_12    1200001  1310000
WP AAAT42063_13    1300001  1410000
WP AAAT42063_14    1400001  1510000
WP AAAT42063_15    1500001  1610000
WP AAAT42063_16    1600001  1710000
WP AAAT42063_17    1700001  1810000
WP AAAT42063_18    1800001  1830121
```

Query Match 11.64; Score 160.4; DB 2; Length 110000;  
Best Local Similarity 49.64; Pred. No. 2.1e-34;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy      1 TAAAGCGTGAACCATGTTAAACAATTTATCACTGCTCATTTAATTTTAAATGATC 60
Db      104439 TAAACATGTAGCTAAATGCGTAAATTTCTGCGCGCTTTTCAAAAAGACAAATGATC 104380
Qy      61 CATGCTAGCAAGCAGGTGACAAAGACAGCCCTTAGCGGTGCTTGAATGCTGCGCAA 120
Db      104379 GAGACTTGGCAAGCAGGCGACCAATATATACCA-----TATCTCGCTTGGCAA 104332
```

```
Qy      121 TACTGCCGATCAAGTGAACAACCTGATATATTTTATGATGCCAATGTGATCACTTT 180
Db      104331 TGTGGGCGTAAATAATTTATATGCTTGTTCATTTGATGAACA-----ATACTTTG 104278
Qy      181 GAGTAAATCTTGTCTAATTAATCGGGGCATCTTGACCGATTAATATACATATCACCA 240
Db      104277 CGATGAAAATTTTGAAGAGGCGCACATCTCGACCAAAACATTAATATATGTTGG 104218
Qy      241 CTGATTAACATATGCTGACAGCTCAATAAATTTTAACTTTTACCTTTACCAAGCCCAAAAT 300
Db      104217 TTGATTAATTAATTAATGCTAAATTTCTGAAAATTAAGCCCTTTTCGTTCCGCTTGCAA 104158
Qy      301 CAATGCCAGTATGATATGCCATTAACCGTGCACAGCCCATCAATGCTGCAATGTTGA 360
Db      104157 CAATATGCAATTTTACCTTCAATAT-----AAGCCAGCCCAATGCGAACAATGTAAT 104107
Qy      361 GCGGATGTTTGTATCCTTTTGAATCATTAATAATATGCTTGTATCAATATCATTAATATA 420
Db      104106 CCCCATTTTGTGCTTTTGAAGTCATTAATCCAAAGAAAGCCATGAGCTTGAATGCACTAA 104047
Qy      421 TTACAGCGATGCGGTAGCCCTTTGAAGTCTTAAAGGTATCAAGATGCGATCCATGGG 480
Db      104046 TTGAAAACGATGATCTAACCTTTGAAATGACGAAAGTGGGTACGAATTTGAATCTAAAT 103987
Qy      481 AAGATTGGCAGCTGCTTAAAGGCAAGTGCAGATAAGCAATGATGATGATGATGATGATGAT 540
Db      103986 AATACCTATAGCTTGTGCGCAATGCTGTGTGCGCAAAATGTTTCAATTAATTTGGGACC 103927
Qy      541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db      103926 AACCAATGATGCTTCTTCAAGGTAAATATCACTTATCTTTTACCATTAATATTTGCTT 103867
Qy      601 ATACCATATCAATAATTAATAAATCTGC-CGTTTGGGTGCTTAAGTAACTTATTTGTTG 659
Db      103866 GCCATTTTCAAGTTTGTAGCCAAATATCCGACATATTTTCGCAAAAGAAACGGTATGTT 103807
Qy      660 TGAATTAATGATGATTTTGGGCAACGCTGTGATGATGATGATGATGATGATGATGATGAT 719
Db      103806 CGCTTGAATTTTGTGTTTCCCAAAAGTCAGCC----- 103775
Qy      720 TGGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db      103774 ---TATCTTCAATTTGTCACACACACTTATGATGATGATGATGATGATGATGATGATGATGAT 103718
Qy      780 CCAATTAATTCACAGATGCTGCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 839
Db      103717 GGGGATTAATCTTCTAATCCATATAGCATGATGATGATGATGATGATGATGATGATGATGAT 103658
Qy      840 TTGCCCCCTTGAAGTGCATAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
Db      103657 TCGCTGCGGAGCTTTTAAAGCTATAGTGTCTCAAGCTGAAAAGCTCTAGTA 103598
Qy      900 CCACCAATTCATATTTTGGATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 959
Db      103597 CATTAAGTTCACATCTTCAATTC---AACATGACAAAGCGGGAATCCCAATATTTCCGC 103541
Qy      960 CGACACCAACATCATGCTGCTGATGTTTGGCATCTGCTGCTACTATATGCTGTAACGTCG 1019
Db      103540 CCAATACCACTTTCACACAGCAGCTTTCGCCATTTCAATPAACATTAAGTATGCGTAC 103481
Qy      1020 TTTTGGCATTTGAGCTGTGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
Db      103480 TTTTACCATTTTGAACCTGTATATCCCAATATGCTTGTGCTGCTGCGGCA----- 103429
Qy      1080 GTGCTTTGAGTGTATCAATGAACAATGGAACATGCTGATCAACGAATGCTTGAAGCTT 1139
Db      103428 -----GAAATTAATTCATATATGCGGATTTACTTCCACTCCGCTTTAA 103388
Qy      1140 TGGCGGCGACAACTACTGCGCTTCTGGGTTAATGCCGGGCTGATGATGATGATGATGATGATGATGAT 1199
Db      103387 GTGCGGTTTGAATTTCTGTTGTTTAAACGCAAGCCTTGGCTTAATTAACATATATATCCG 103328
Qy      1200 CTGAGCTTAAGATTCATGCTCCAAAGCTGCCAAATGATGAATGAACGCTGTGGGAGTT 1259
```

Db 103327 TTTCAGTAACATTCCTGATTTAACTACAGATATAGAGGGAATTTTGAAGAACTT 103268  
Qy 1260 TATCTGCCAAGTGTGATTTGGATTTACATCGATGATGCTTAATCTTTGTGACCGTGATTA 1319  
Db 103267 TATCAATACCAAGTGTGATTTTTCGAGTATCAATCAACGAAATATTTGCGTTGGGATA 103208  
Qy 1320 CCAAAAAATTCACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 1377  
Db 103207 AGAGATATATCAACAAAGAAAGACCTGTTTGGCCAGCCGAGTATGATATATTTT 103150

## RESULT 12

AA142063\_12/c

Continuation (13 of 19) of AA142063 from base 120001 (Haemophilus influenzae complete g

Sequence Split into 19 fragments LOCUS AA142063 Accession AA142063

| WP | Fragment Name | Begin   | End     |
|----|---------------|---------|---------|
| WP | AA142063_00   | 1       | 110000  |
| WP | AA142063_01   | 100001  | 210000  |
| WP | AA142063_02   | 200001  | 310000  |
| WP | AA142063_03   | 300001  | 410000  |
| WP | AA142063_04   | 400001  | 510000  |
| WP | AA142063_05   | 500001  | 610000  |
| WP | AA142063_06   | 600001  | 710000  |
| WP | AA142063_07   | 700001  | 810000  |
| WP | AA142063_08   | 800001  | 910000  |
| WP | AA142063_09   | 900001  | 1010000 |
| WP | AA142063_10   | 1000001 | 1110000 |
| WP | AA142063_11   | 1100001 | 1210000 |
| WP | AA142063_12   | 1200001 | 1310000 |
| WP | AA142063_13   | 1300001 | 1410000 |
| WP | AA142063_14   | 1400001 | 1510000 |
| WP | AA142063_15   | 1500001 | 1610000 |
| WP | AA142063_16   | 1600001 | 1710000 |
| WP | AA142063_17   | 1700001 | 1810000 |
| WP | AA142063_18   | 1800001 | 1850121 |

Query Match 11.6%; Score 160.4; DB 2; Length 110000;

Best Local Similarity 49.6%; Pred. No. 2,1e-34;

Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy 1 TAAAGCTGAACCATGTTAAACATTAATCACTGCTGCTATTAATTTTAAATGATC 60  
Db 4439 TAAACATTTAGAGCTTAATGCGTAATCTTTCGCGCTTTTCAAAAAGAACATGATC 4380  
Qy 61 CAGCTGACCAACAGCTGACCAACAGCTGACCAACAGCTGACCTGCTGCTGCTGCA 120  
Db 4379 GAGACTTGACCAACAGCTGACCAACAGCTGACCTGCTGCTGCTGCTGCTGCA 4332  
Qy 121 TACTGCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAATGTCATCCTT 180  
Db 4331 TGTGGGCGTAAATTAATCTATCGCTTGTTCATGTAATGCAACA-----ATACTTG 4278  
Qy 181 GAGTAAATCTTGCTCAATTAATGCGGCACTTGAACCAATTAATACACATTCACCA 240  
Db 4277 CGATGAAAAATTTTGAACAGCGCACATCTCGACAAAACAAATAAATATGTTGG 4218  
Qy 241 CTGATTAACATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCCAACCAAT 300  
Db 4217 TTGATTAATTAATCACTAATTTGAAAAATCAGCCCTTTTCGCTCTCGCTACCA 4158  
Qy 301 CAATGCAATGATATATGCAATTAACCGTGCACAGCCCAATCAATTTCTGCAATGTTGA 360  
Db 4157 CAATGCAATTTTACCCCAATATA-----AAGCCCAAGCAATGCAAGAACTGTAAT 4107  
Qy 361 GCGGATTTTGTACCTTTTGAATCAATTAATATGCTTGTGATCAATATCAATATA 420  
Db 4106 CCCCACTTTTGTCTTTTGAAGTATTAATCAACGAATGCAATTAAGCTTGAATCACTAA 4047  
Qy 421 TTCAACAGGATGCGGTGCGCTTTGAAAGCTTAAAGGTATCAACAGCATGCGATCAGGG 480  
Db 4046 TTGAAAAAGATATTAACCTTTGAAATGACGAAGTGGGTACGAATGATTAATTT 3987

Qy 481 AAGATTGAGCTGTGCTTAAGCAAGTGCAGATTAAGCATTAAGATTGATGCTTGC 540  
Db 3986 AATACATATAGCTTGTGCGCAATGCTGTGTGCAAAAATGTTATTAATTAATGAGGAC 3927  
Qy 541 TTTGATTTTAAATGATGATTTGTAAGTTGTTTGTGATGATTAATGCCATACC 600  
Db 3926 AACCAATGATGCTTCTTCAAGGTAAATCACTTCATCTTTTACCATTAATATGCTT 3867  
Qy 601 ATCAGCATCAATTAATTAATTAATCTGC-CGTTGTGCGCTAAGTAAATGTTTG 659  
Db 3866 GCCATTTTCACTTTTAAAGCAATTAATCGACTATTTTCCGCAAAAAGAAAGGTATGTT 3807  
Qy 660 TGATTAATGATGATTTGGGCAACGCTGTGATGATGATGATGATGATGATGATGAT 719  
Db 3806 CGCTTATTTTGTGTTTCCCAAAAGTCAGCC----- 3775  
Qy 720 TGGCATGCTGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 779  
Db 3774 ---TATCTTCATTTGTTCAACACACTTATGATGATGATGATGATGATGATGATGAT 3718  
Qy 780 CCAATTAATCAACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 839  
Db 3717 GCGGATTAATTTTAAATCAATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3658  
Qy 840 TTGCGCTTGTGCTGCAATTTAGATGATGATGATGATGATGATGATGATGATGATGAT 899  
Db 3657 TCGCTGCGGAGCTTTTAAAGCTAATGATGATGATGATGATGATGATGATGATGATGAT 3598  
Qy 900 CCAACCAATCAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 959  
Db 3597 CATTAATGCTCAATCTTCAATTC--AACATTAACAAAGGGAATCCCAATATTTCCG 3541  
Qy 960 CGACCAACCAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019  
Db 3540 CCAATCAATTTTCAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3481  
Qy 1020 TTTTGGATTTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
Db 3480 TTTTACATTTTGAACCTGTAATCCCAATGATGATGATGATGATGATGATGATGATGAT 3429  
Qy 1080 GTGCTTGTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1139  
Db 3428 -----GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3388  
Qy 1140 TGGCGGCAATATATGCGGCTTCTGCTGATTAATGCGGCGGCTGATGATGATGATGAT 1199  
Db 3387 GTGCGTTGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3328  
Qy 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259  
Db 3327 TTTCAGTAAACATTTCTGATTTTAACTACAGTATGAAGGGAATTTTGAAGAAATT 3268  
Qy 1260 TATCTGCCAAGTGTGATTTGGATTTACATCGATGATGATGATGATGATGATGATGAT 1319  
Db 3267 TATCAATACCAAGTGTGATTTTTCGATGATCAATCAACGAATTAATGCTGTTGGATA 3208  
Qy 1320 CCAAAAAATTCACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGATTTT 1377  
Db 3207 AGAGATATATCAACAAAGAAAGACCTGTTTGCAGAGCCGAGTATGATATATTTT 3150

## RESULT 13

AEB39175\_30

Continuation (31 of 35) of AEB39175 from base 300001 (L. pneumophila DNA SEQ ID NO 3507.

Sequence Split into 35 fragments LOCUS AEB39175 Accession Aeb39175

| WP | Fragment Name | Begin  | End    |
|----|---------------|--------|--------|
| WP | AEB39175_00   | 1      | 110000 |
| WP | AEB39175_01   | 100001 | 210000 |
| WP | AEB39175_02   | 200001 | 310000 |
| WP | AEB39175_03   | 300001 | 410000 |
| WP | AEB39175_04   | 400001 | 510000 |
| WP | AEB39175_05   | 500001 | 610000 |
| WP | AEB39175_06   | 600001 | 710000 |

WP AEB39175\_07 700001 810000  
WP AEB39175\_08 800001 910000  
WP AEB39175\_09 900001 1010000  
WP AEB39175\_10 1000001 1110000  
WP AEB39175\_11 1100001 1210000  
WP AEB39175\_12 1200001 1310000  
WP AEB39175\_13 1300001 1410000  
WP AEB39175\_14 1400001 1510000  
WP AEB39175\_15 1500001 1610000  
WP AEB39175\_16 1600001 1710000  
WP AEB39175\_17 1700001 1810000  
WP AEB39175\_18 1800001 1910000  
WP AEB39175\_19 1900001 2010000  
WP AEB39175\_20 2000001 2110000  
WP AEB39175\_21 2100001 2210000  
WP AEB39175\_22 2200001 2310000  
WP AEB39175\_23 2300001 2410000  
WP AEB39175\_24 2400001 2510000  
WP AEB39175\_25 2500001 2610000  
WP AEB39175\_26 2600001 2710000  
WP AEB39175\_27 2700001 2810000  
WP AEB39175\_28 2800001 2910000  
WP AEB39175\_29 2900001 3010000  
WP AEB39175\_30 3000001 3110000  
WP AEB39175\_31 3100001 3210000  
WP AEB39175\_32 3200001 3310000  
WP AEB39175\_33 3300001 3410000  
WP AEB39175\_34 3400001 3503610

Query Match 11.4%; Score 157.6; DB 14; Length 110000;  
Best Local Similarity 51.9%; Pred. No. 1.4e-33;  
Matches 467; Conservativity 0; Mismatches 409; Indels 24; Gaps 4;

QY 170 GCATCAGCTTTGAGTAATCTTGTCAATTATCGGGGACCTTGCAGGATTAATACACA 229  
DB 39768 GGAACCTACTTGGCTAAAGCCGATTCATCTTGCAGCATCTTCACCAATCAAAACATG 39827  
QY 230 CTATCCACCACTGCATACATATGTCAGAGCTCACTAAATCTTACCTTACCAAGC 289  
DB 39828 GAAAGCAAAATCTGAAACCTGGGCTGATTAATCTTGAATAATCGCCCTTTTCCCTGT 39887  
QY 290 CCAACCAAAATCAATGCGATGATATGCAATAAACCGTCCAGCCCATCAATTTCT 349  
DB 39888 CCAACCGCAATTAAGAACATTTTCCCTTGCATAGAAC--TCTATTTCAATTAATGCA 39944  
QY 350 GCATGCTTACGCGGATTTGTAACCTTTGATCAATTAATAATGCTTGTCTATCAATA 409  
DB 39945 GAGATGTGACACCAATATTTGTTCTTGAATCAATTAATCCAACTTCCATCCACT 40004  
QY 410 TCATCATATATTCACAGGATGCGGTAGCCCTTGAAGTCTTAAGGTATCAAGCATG 469  
DB 40005 TCTCTTAACCACTGCGATGATGCGAGTCCAGGAGGTTTAAACATTAAGATG 40064  
QY 470 GCATCCATGGAAGATGAGAGCTGTGCTTAAGGCAATGACATTAAGCATTAAGAG 529  
DB 40065 TGTGATTAAGATTCCTCGCTGCTGCTTAAGGCAATGACATTAAGCATTAAGCA 40124  
QY 530 TTATGCTTGGCTTTGATTTTAAATGATGATGTAAGATTGTTTGTGATGATAT 589  
DB 40125 TTATGACACCCCTTTTAAATGATGATGATGATGATGATGATGATGATGATGATG 40184  
QY 590 AATGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 649  
DB 40185 AATTAAGTATGATTTCTTTGTTCAATGACCCCAATTTCCATTT-----GAG 40233  
QY 650 CTATTTGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 709  
DB 40234 GTGCGTCTTGGCAAAAGATGATGATTTTATCTGTTTGAACGATGATGATGATGATG 40293  
QY 710 TTGGCTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 769  
DB 40294 T-----ATGAGGCTCTCTCGATTAATTAAGACGCTTGGACCAAGATTAATCTCT 40346

QY 770 AATTTTGTGCCAATATATCCATGCGATGCGGATCAAGATGATGACGGGAGAGA 829  
DB 40347 TGTTTTCTGCTGCTATAGGCTTCCATAGTGTATGCTGTCCAAATGATGAGCGTTACA 40406  
QY 830 TTTAAATTTGTTCCCTTGTAGAGTGCCTCAATTTGATGATGCTCAAGCTGAATCTTGAC 889  
DB 40407 TTAAGATGCTGCCAACCACTGCGCAAGAAATTAAGTTAAATTCATTAATAAATCTGAT 40466  
QY 890 AGCTCAAGCAGCAACCAATC---CATATTTGATGCTTAATTAATCAATGAGAGGTG 946  
DB 40467 AACTCCAAATGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 40526  
QY 947 CCAATATTAACCGCGACACCAACATCATCTGATGTTTGTCCATCTGCGCTACTAAT 1006  
DB 40527 CCATATTTTCCGCTATGAGTACGATGCGCAACCGCGCTTACCATTTCTCCACCAAA 40586  
QY 1007 GTGCTAACAGTGTGTTGGATTTGAGCTGTGATGAGCGATGATGATGATGATGATG 1066  
DB 40587 GTGCTAACAGTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 40646

RESULT 14  
AEB42401\_28  
Continuation (29 of 34) of AEB42401 from base 2800001 (L. pneumonia DNA SEQ ID NO 6733  
WP Sequence Split into 34 fragments LOCUS AEB42401 Accession Aeb42401

WP Fragment Name Begin End  
WP AEB42401\_00 1 110000  
WP AEB42401\_01 100001 210000  
WP AEB42401\_02 200001 310000  
WP AEB42401\_03 300001 410000  
WP AEB42401\_04 400001 510000  
WP AEB42401\_05 500001 610000  
WP AEB42401\_06 600001 710000  
WP AEB42401\_07 700001 810000  
WP AEB42401\_08 800001 910000  
WP AEB42401\_09 900001 1010000  
WP AEB42401\_10 1000001 1110000  
WP AEB42401\_11 1100001 1210000  
WP AEB42401\_12 1200001 1310000  
WP AEB42401\_13 1300001 1410000  
WP AEB42401\_14 1400001 1510000  
WP AEB42401\_15 1500001 1610000  
WP AEB42401\_16 1600001 1710000  
WP AEB42401\_17 1700001 1810000  
WP AEB42401\_18 1800001 1910000  
WP AEB42401\_19 1900001 2010000  
WP AEB42401\_20 2000001 2110000  
WP AEB42401\_21 2100001 2210000  
WP AEB42401\_22 2200001 2310000  
WP AEB42401\_23 2300001 2410000  
WP AEB42401\_24 2400001 2510000  
WP AEB42401\_25 2500001 2610000  
WP AEB42401\_26 2600001 2710000  
WP AEB42401\_27 2700001 2810000  
WP AEB42401\_28 2800001 2910000  
WP AEB42401\_29 2900001 3010000  
WP AEB42401\_30 3000001 3110000  
WP AEB42401\_31 3100001 3210000  
WP AEB42401\_32 3200001 3310000  
WP AEB42401\_33 3300001 3345687

Query Match 11.4%; Score 157.6; DB 14; Length 110000;  
Best Local Similarity 51.9%; Pred. No. 1.4e-33;  
Matches 467; Conservativity 0; Mismatches 409; Indels 24; Gaps 4;

QY 170 GCATCAGCTTTGAGTAATCTTGTCAATTATCGGGGACCTTGCAGGATTAATACACA 229  
DB 96841 GGAACCTACTTGGCTAAAGCCGATTCATCTTGCAGCATCTTCACCAATCAAAACATG 96900  
QY 230 CTATCCACCACTGCATACATATGTCAGAGCTCACTAAATCTTACCTTACCAAGC 289  
DB 96901 GAAAGCAAAATCTGAAACCTGGGCTTAATTTGGAATAATCCGCCCTTTTCCCTGT 96960







```
Db      5901 CAAAACAGTCGACGCGCAGCTTTTAAGTATTAAGTTGTCGAAGTGAATAAG 5960
QY      892 CTCAGACACCAACCAATCCATATTTGGATCGTTAATTAATTCAGAGCGCGTCCAT 951
Db      5961 CTCATAGTAACTAAAGTTGACAAATCTTCATTC---AACAAAGACAAAGCGGAAATCCCAAT 6017
QY      952 ATTACGCGCGACACCAACATCATGCGTGATGTTTGGCCATCTCGGCTACTAATGTGT 1011
Db      6018 ATTTCGCGCCATACCAACTTTCACGCGACAGAGCTTTCGCAATTCATTAAGTAAT 6077
QY      1012 AACAGTGCCTTTGGCATTTGAGCGTGTGATGCGCATGATGATGTTGTGCTTGAGT 1071
Db      6078 AACAGTACTTTTTCATTTGATTAACCTGTATCCCAATATGGCTTTGTGCTGCGCG--- 6134
QY      1072 TAGATCTCGTCTTTGAGTATCAATGAACATTTGAACATCGCTGATCAAGGAATGCC 1131
Db      6135 -----GCAGAAATAATTCATATTCGCCGATTAATCTTCACACTCC 6170
QY      1132 TTGAGCTTTGGCGGCGACAAATTAATCTCGGCGTTCTTGGTTAATGCGGCGCTGATGATGAT 1191
Db      6171 CGCTTTAAGTCGCGTTTGAATTTCTGTGTGTTTACCGCAAGCCCTGGGCTAATTAACAT 6230
QY      1192 CTGATCGGCTGAGCTTAAGAGTTCATGTCCAAGCTGCCAAATGACAAATGAACGCCCTGT 1251
Db      6231 CATATCGGCTTCAAGTAACTTCTGATTTAACTACAGTATGAAGAGGATATTTTG 6290
QY      1252 GGGCAGTTTATCTGCCAAGTAAAGTGGGATTACATGATGATGCTAATTTGTGACC 1311
Db      6291 AGGAAGTTTATCAATACCAAGTAGGCTTTTTCGGGTATCAATCACAGAAATATTAGCCTG 6350
QY      1312 GTGATTGACCAAAAATTTCACAGCTGACGTCCAGAGCTGCCCAACCGACGAGCAT 1371
Db      6351 TTGGATTAAGAGTAATTCACACAGAAAGACCTGTTTTCGCAAGCCCAATGATCGTAAT 6410
QY      1372 TTTTTT 1377
Db      6411 ATTTT 6416
```

Search completed: May 18, 2006, 13:52:28  
Job time : 955 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2006, 08:54:03 ; Search time 6244 Seconds  
(without alignments)  
12358.853 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736  
Perfect score: 1380  
Sequence: 1 taagcgtagacatgtaaa.....acgacgcatatttttga 1380

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96455510

Minimum DB seq length: 22

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:\*  
1: gb\_esc1:\*  
2: gb\_esc3:\*  
3: gb\_esc4:\*  
4: gb\_esc5:\*  
5: gb\_esc6:\*  
6: gb\_esc7:\*  
7: gb\_esc8:\*  
8: gb\_esc9:\*  
9: gb\_esc10:\*  
10: gb\_esc11:\*  
11: gb\_esc12:\*  
12: gb\_esc13:\*  
13: gb\_esc14:\*  
14: gb\_esc15:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| C 1        | 98.2  | 7.1         | 817    | 13    | CZ543630 SRAA-aads  |
| C 2        | 92.8  | 6.7         | 709    | 5     | CF306940 HDAL--05-  |
| C 3        | 92    | 6.7         | 892    | 12    | BZ579116 meht.6205  |
| C 4        | 80.8  | 5.9         | 916    | 14    | DUT73587 APK13331.  |
| C 5        | 80.2  | 5.8         | 1025   | 14    | DUT773354 APK3419.  |
| C 6        | 77.2  | 5.6         | 555    | 14    | AG943479 Drosophila |
| C 7        | 76.4  | 5.5         | 977    | 14    | DUT48329 ASNC465.b  |
| C 8        | 69.8  | 5.1         | 1445   | 12    | BZ568959 pac82-164  |
| C 9        | 62.2  | 4.5         | 623    | 14    | AG952088 Drosophila |
| C 10       | 59.6  | 4.3         | 928    | 14    | DUT69542 APK1505.   |
| C 11       | 56.8  | 4.1         | 628    | 13    | CL691264 PRI0156a   |
| C 12       | 55    | 4.0         | 474    | 9     | DR032465 ba020270   |
| C 13       | 54.2  | 3.9         | 1024   | 14    | DUT732353 APK1690.  |
| C 14       | 52.6  | 3.8         | 1171   | 11    | BZ559806 pac82-164  |
| C 15       | 50.2  | 3.6         | 838    | 13    | CZ545153 SRAA-aad6  |
| C 16       | 49.8  | 3.6         | 572    | 11    | BH795426 BMBAC376G  |
| C 17       | 49.4  | 3.6         | 452    | 7     | BE184674 PM4-HT068  |
| C 18       | 48.6  | 3.5         | 958    | 14    | DUT71470 APK2473.   |
| C 19       | 47.6  | 3.4         | 1101   | 14    | CNS00100 Drosophila |

|      |      |     |      |    |          |                     |
|------|------|-----|------|----|----------|---------------------|
| C 20 | 47.2 | 3.4 | 1101 | 14 | CNS002MK | AL097862 Drosophila |
| C 21 | 46   | 3.3 | 823  | 10 | DR505933 | MS02715.B           |
| C 22 | 45.4 | 3.3 | 856  | 14 | AJ863943 | Palaeontia          |
| C 23 | 45   | 3.3 | 590  | 1  | AU287701 | AU287701            |
| C 24 | 45   | 3.3 | 757  | 9  | DR473894 | DR473894 WS00960.B  |
| C 25 | 45   | 3.3 | 907  | 10 | DN922769 | DN922769 43863.2.C  |
| C 26 | 44.4 | 3.2 | 442  | 10 | DR565020 | DR565020 WS02632.C  |
| C 27 | 44.4 | 3.2 | 490  | 8  | CO253110 | CO253110 WS00818.B  |
| C 28 | 44.4 | 3.2 | 794  | 10 | DR472810 | DR472810 WS00960.B  |
| C 29 | 44   | 3.2 | 918  | 14 | DUT72406 | APK29941.           |
| C 30 | 43.6 | 3.2 | 692  | 11 | BH983612 | BH983612 ode20g11.  |
| C 31 | 43   | 3.1 | 1035 | 14 | CNS002D5 | AL097523 Drosophila |
| C 32 | 42.8 | 3.1 | 1101 | 14 | CNS0100X | AL098379 Drosophila |
| C 33 | 42.4 | 3.1 | 934  | 14 | DUT91861 | DUT91861 APK4300.   |
| C 34 | 42.4 | 3.1 | 1101 | 14 | CNS0039G | AL063921 Drosophila |
| C 35 | 42.2 | 3.1 | 747  | 13 | CZ668937 | OM_Ba022            |
| C 36 | 42.2 | 3.1 | 915  | 11 | AZ543308 | AZ543308 ENTG93TR   |
| C 37 | 42   | 3.0 | 941  | 14 | CNS05L54 | AL342337 Tetradon   |
| C 38 | 42   | 3.0 | 1022 | 14 | CNS07ANB | AL36797 T3 end of   |
| C 39 | 41.8 | 3.0 | 517  | 1  | AU287702 | AU287702            |
| C 40 | 41.8 | 3.0 | 1101 | 14 | CNS000D1 | AL065414 Drosophila |
| C 41 | 41.6 | 3.0 | 616  | 11 | BH383040 | BH383040 AG-ND-127  |
| C 42 | 41.6 | 3.0 | 714  | 11 | BH374901 | BH374901 AG-ND-127  |
| C 43 | 41.6 | 3.0 | 747  | 4  | CA920035 | CA920035 EST631753  |
| C 44 | 41.4 | 3.0 | 639  | 9  | CX639436 | CX639436 UCRPT02.4  |
| C 45 | 41.4 | 3.0 | 760  | 11 | BH367641 | BH367641 AG-ND-127  |

#### ALIGNMENTS

RESULT 1  
CZ543630/817 bp DNA linear GSS 13-MAY-2005  
LOCUS SRAA-aad51b03.g1 Strongyloides ratii whole genome shotgun library  
DEFINITION (SRAAGS 004) Strongyloides ratii genomic, genomic survey sequence.  
ACCESSION CZ543630  
KEYWORDS GSS.  
SOURCE Strongyloides ratii  
ORGANISM Strongyloides ratii  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Panagrolaimoidea; Strongyloidea; Strongyloidea.  
REFERENCE 1 (bases 1 to 817)  
Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Page,D.,  
Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,K.H.,  
Clifton,S.W. and Wilson,R.  
Genome Survey sequences from the rat parasitic nematode  
Strongyloides ratii  
Published (2005)  
JOURNAL Contact: Mitreva M  
COMMENT Washington University in St. Louis  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: nematode@wustl.wustl.edu  
Genomic DNA was provided by Fiona Thompson  
(F.Thompson@bristol.ac.uk) and Mark Viney  
(Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,  
UK.  
Class: shotgun.  
Location/Qualifiers  
1..817  
/organism="Strongyloides ratii"  
/mol\_type="genomic DNA"  
/strain="Isofemale line ED321 heterogenic"  
/db\_xref="taxon:34506"  
/dev\_stage="infective larval stage (iL3)"  
/lab\_host="GSI0"  
/clone\_lib="Strongyloides ratii whole genome shotgun  
library (SRAAGS 004)"  
/note="Vector: POTW13; Site\_1: BstXI; Site\_2: BstXI;





```

source
1. .1025
/organism="uncultured marine microorganism
HF4000.12-21-03"
/mol_type="genomic DNA"
/db_xref="taxon:361148"
/clone="HF4000_1384108F16"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"
/clone_1lb="HF4000.12-21-03"
/notes="Vector: pCCLFOS; North Pacific Subtropical Gyre
(Hawaii) picoplankton genomic fosmid DNA library prepared
from marine picoplankton in the less than 1.6 um, greater
than 0.22 um fraction. Picoplankton collected at 4000 m
depth on 10/7/2002, Coordinates: 22.45 N, 158 W. Sample
Date: 12/21/2003 Coordinates: 22.45 N, 158 W Depth 4000 m
Temperature: 1.46 C Salinity: 34.69 Oxygen: 147.8 umol/kg"

ORIGIN

Query Match 5.8%; Score 80.2; DB 14; Length 1025;
Best Local Similarity 56.0%; Pred. No. 2.6e-11;
Matches 173; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 762 AATACGCAATTTTGTGCCAATATCCACCATGCTGCGCCATCAAGATGATCAG 821
Db AGATGGCGCTCTGGCCGCGCAGGAGTCACTGAAGCCGTTGAGCGATCCAGATGCTCT 133
QY 822 CGAGAGATTTAAATGTTGCCCTTGAGCTGCGCAATTTGAGATATGCTCAAGCTGAA 881
Db 134 GGAACACATTGAGGATGTGGCCACCTGGCATTCAGGCTGTAGTGTCTTCCAATGGA 193
QY 882 AACTTGACAGCTCAAGCAGCACCACCAATTCATATTTGATCGTTAATATTCAGTGAC 941
Db 194 AGCTGACAGCTCCAGAGACGTACAGACC---GGCGCTGCTGGCAGAGATCCAGCCGCG 250
QY 942 GCGTCCCAATTTTACCGCCGACACCAACATCATGCTGATGTTTCCATCTCGCTTA 1001
Db 251 GGAATCCCAAGTTGCGCCCAACCCCGCGCTGACGCGCCCAAGCTTCGCCCATTCACCGA 310
QY 1002 CTATATGCTGAACAGTCTTTTGATTTGAGCCTGTGATGGCGATATTTGTGTGTTG 1061
Db 311 GCAGGCTGTGACGCTCTCTTGCGATTGGAACCGGTGATCGCCACGAGGGCGGCTGG 370
QY 1062 TTGCTTGAG 1070
Db 371 CGGCACGCG 379

RESULT 6
AG943479 555 bp DNA linear GSS 01-FEB-2005
LOCUS Drosophila ananassae DNA, clone: DNBI-017M23.R.fa, genomic survey
DEFINITION
ACCESSION AG943479
VERSION AG943479.1 GI:58464900
KEYWORDS GSS.
SOURCE Drosophila ananassae
ORGANISM Drosophila ananassae
Bukaryota: Metazoa: Arthropoda: Hexapoda: Insecta: Pterygota;
Neoptera: Endopterygota; Diptera: Brachycera; Muscomorpha;
Phylozoa: Drosophilidae; Drosophila.
1
Hattori, M., Toyoda, A., Murakami, K., Kuroki, Y., Fujiyama, A.,
Toshio, T. K. and Sakaki, Y.
BAC end sequences of library DNBI
Unpublished
2 (bases 1 to 555)
Hattori, M.
Direct Submission
Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hsp.gsc.riken.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

```

```

COMMENT
Clones are derived from the BAC library DNBI
For BAC library availability, please contact Mase-Toshi Yamamoto
(Yamamoto@itc.jp).
Submitted (30-11-2004) by Masahira Hattori,
RIKEN, Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, Tel: 81-45-503-9111,
Fax: 81-45-503-9170)
This work was done in collaboration with Yamamoto, M-T. Drosophila
Genetic Resource Center
Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
Tel: 81-75-873-2660 FAX: 81-75-861-0881
PRIMERS
Sequencing : R
LIBRARY
Vector : pKS150
R.site 1 : SacI
R.site 2 : SacI
Location/Qualifiers
1. .555
/organism="Drosophila ananassae"
/mol_type="genomic DNA"
/db_xref="taxon:7217"
/clone="DNBI-017M23.R.fa"
/clone_1lb="DNBI Drosophila BAC library"

ORIGIN

Query Match 5.6%; Score 77.2; DB 14; Length 555;
Best Local Similarity 53.3%; Pred. No. 1.6e-10;
Matches 163; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 758 TCATAAATACGAATTTTGTGCCAATATCCACCATGCTGCGCATCAAGTGA 817
Db 426 TCATAAATATGATGTTTGTGCTGATATATCCGTGATATTCACGAGAGCTCTAATGA 367
QY 818 TCAGCGAAGATTTAAATTTGTCCTTGAAGTCCAAATTTAGATATGCTCAAC 877
Db 366 TCTGGAATATATTAATGAACATGAATATACGCTGAATATTCAGAAAGTTCTAAC 307
QY 878 TGAATCTTGACGCTCAAGCACCACCAATCATATTTGATGCTTAATATCACT 937
Db 306 ATATTAAGAGACATCTCTTGAACATATACGCTTATATTCATTAAGAAAGACCAT 247
QY 938 GCAGCGTGCATATATTCGCGCAGACCAACATCATGCTGATGTTTTCATCTCG 997
Db 246 GCAGCAGACCTAATATCTCTCTGCGATCGCAGAAATCAAGCTTCATTTAAAGATCT 187
QY 998 CTAATATGCTGAACAGTCTTTTGCAATTTGAGCCTGTGATGCGATGATGTGTG 1057
Db 186 GTTAATTAACAGCTGTGTGATTTCCATTTGTGCTGTGATGCAACAAACGTGCT 127
QY 1058 TTGTT 1063
Db 126 TTGTAT 121

RESULT 7
DUT48329 977 bp DNA linear GSS 27-JAN-2006
LOCUS ASNC65.b2 HF10_10-07-02 uncultured marine microorganism
DEFINITION
ACCESSION DUT48329
VERSION DUT48329.1 GI:85758165
KEYWORDS GSS.
SOURCE uncultured marine microorganism HF10_10-07-02
ORGANISM uncultured marine microorganism HF10_10-07-02
unclassified sequences; environmental samples.
1 (bases 1 to 977)
Delong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,
Frigard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W.
and Karl, D.M.
Comparative genomics reveals ecological trends in stratified
microbial communities in the ocean's interior

```

JOURNAL  
COMMENT

Science (2006) In press  
Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,  
Kerrie Barry, Tijana Glavinadelirio, David Bruce, Paul Richardson  
and Edward Delong  
US DOE Joint Genome Institute  
US DOE Joint Genome Institute  
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA  
Tel: 617-253-5271  
Fax: 617-253-2679  
Email: P.Richardson@lbl.gov; delong@mit.edu  
North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid  
DNA library prepared from marine picoplankton in the less than 1.6  
um, greater than 0.22 um fraction. Sample Date: 10/7/2002  
Coordinates: 22.45 N, 158 W Depth: 10 m Temperature: 26.4 C  
Salinity 35.08 psu Oxygen: 204.6 umol/kg  
Class: fosmid ends.

FEATURES  
source

1. 977  
/organism="uncultured marine microorganism HF10\_10-07-02"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:361145"  
/clone="HF010\_001A11"  
/cell\_type="marine picoplankton, less than 1.8 um, greater  
than 0.22 um fraction"  
/clone\_lib="HF10\_10-07-02"  
/note="Vector: pC1FOS; North Pacific Subtropical Gyre  
(Hawaii) picoplankton genomic fosmid DNA library prepared  
from marine picoplankton in the less than 1.6 um, greater  
than 0.22 um fraction. Picoplankton collected at 10 m  
depth on 10/7/2002, Coordinates: 22.45 N, 158 W. Sample  
Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth: 10 m  
Temperature: 26.4 C Salinity 35.08 psu Oxygen: 204.6  
umol/kg"

## ORIGIN

Query Match 5.5%; Score 76.4; DB 14; Length 977;  
Best Local Similarity 49.5%; Pred. No. 2-10;  
Matches 271; Conservative 0; Mismatches 256; Indels 21; Gaps 2;

QY 338 CCATCAATTCCTGCAATGTTGAGCCGATGTTGTACCTTTGAAATTAATATGACC 397  
DB 346 CTACGAACACCTGATTAAGTGGCCCCCTCACTAGTTCCTTGGAAATCGTTGATATAGTT 405  
QY 398 TTGCTATCAATATCATATATATATACAGGATGCGGTACCTTTGAAAGTCTTAAG 457  
DB 406 ACCGCCGAATGACTGACCTCTTTCACAGGATGAGGCAATCTTGAACCTTTATATA 465  
QY 458 GTATCAAGCATGCGATCCATGGAAGATTGGCAGCTGTGCTTAAGCAAGTCAGATAG 517  
DB 466 GCTGATCCATTTGATCACTGGGATACCTGCCATATCACCAATCGCTATGCTGCCACA 525  
QY 518 GCATGAGTAGTATGCTGCTTGAATTTTAAATGATGATGATGATGATGATGATGATG 577  
DB 526 AAATTCATCAAAATGATGCTCACTTGCAGAGGTATTTCTTCAATGCAACAGTAATCT 585  
QY 578 TTTTGAATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 637  
DB 586 AGACCTGCACTGATTAATGCTTCCCTCAGTCTCTACCAACCGTATGATGATGATG 645  
QY 638 TGGCTAAGTATGATTTGTTGATATGATGATGATGATGATGATGATGATGATGATG 697  
DB 646 GGTGATCAATACCAAACTTTTCAAGTAC-----TTAAGAGCTT 687  
QY 698 GTCAAGCATGCTGGCTAAATGCGATCGTCTTGATGATGATGATGATGATGATGATG 757  
DB 688 GTGTCCAAAGGCTTAAATGCTTAAATGATGATGATGATGATGATGATGATGATG 744  
QY 758 TCAAAATAGCAATTTTGTGCAATATCCACATGATGATGATGATGATGATGATGATG 817  
DB 745 TGGTAAATCTCTGCTTATGCGCTGATTAATGCTGATGATGATGATGATGATGATG 804  
QY 818 TCAAGGAGATTTAAATTTGCTTGAAGCTGCAATTTGATGATGATGATGATGATG 877

DB 805 TCATTTGAATGTTTGAATAGCGGCAACTTTTGATCCAAATTAGAAGTACTTAAC 864  
QY 878 TGAAGACT 885  
DB 865 TGAAGACT 872

## RESULT 8

## LOCUS

BZ568959 1445 bp DNA linear GSS 17-DEC-2002  
pac2-164\_8171.y2 pac2-164 Pseudomonas aeruginosa genomic clone  
pac2-164\_8171, genomic survey sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

JOURNAL  
COMMENT

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
Burns, J.L., Kaul, R., and Olsen, M.V.  
Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) In press  
Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

FEATURES  
source

1. 1445  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pac2-164\_8171"  
/clone\_lib="pac2-164"  
/note="clinical isolate 2-164 Whole genomic shotgun  
library."

## ORIGIN

Query Match 5.1%; Score 69.8; DB 12; Length 1445;  
Best Local Similarity 55.0%; Pred. No. 2,9e-08;  
Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 327 CCGTGCACCAAGCCATTAATGTCGCAATGTTGAGCGGATGTTGTACCTTTGAATCAT 386  
DB 87 CGGACCCACGCCCCCTGATGCGCGCAGGCGGCGGCGGAGCTTGGAATCGT 146  
QY 387 TAAATATGCTTGTATCAATATATATATATATATATATATATATATATATATATAT 446  
DB 147 CGTAGTAGCTCAGCGCCTGCGCTCGGTATCCAGTGGCAGGATGAGCAGGCGGAAA 206  
QY 447 AAGTCTTAAGGATATCAAGATGCGATCCATGGAAGATTGGAGCTGTGCTTAAGCAA 506  
DB 207 ACCCTTCACAGCGCGCAGCATGCGTCGAACGCGAGCCGACCGCATGCGCCAGCCCA 266  
QY 507 GTGCAAGTAAGGATGAGTATGATGATGATGATGATGATGATGATGATGATGATG 566  
DB 267 GCGGCGGAGCGGTTGGAATGATGATGATGATGATGATGATGATGATGATGATG 326  
QY 567 AAGTTTGT 575  
DB 327 GCAAGCTGT 335

## RESULT 9

## LOCUS

AG952088 623 bp DNA linear GSS 01-FEB-2005



```

DEFINITION  Drosophila sechellia DNA, clone: DSEI-004A05.F.fa, genomic survey
ACCESSION  AG952088
VERSION    AG952088.1  GI:58473786
KEYWORDS   GSS.
SOURCE     Drosophila sechellia
            Drosophila sechellia
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1
AUTHORS   Hattori,M., Toyoda,A., Murakami,K., Kuroki,Y., Fujiyama,A.,
            Toshi,T.K. and Sakaki,Y.
TITLE      BAC end sequences of library DNBI
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 623)
AUTHORS   Hattori,M.
TITLE      Direct Submission
JOURNAL    Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Saitama-shi, Tsukuba-shi, Ibaraki, Japan, 305-3858, Japan
            (E-mail:hattori@gs.c.riken.jp, URL:http://hgp.gs.c.riken.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            (E-mail:hattori@gs.c.riken.jp, URL:http://hgp.gs.c.riken.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the BAC library DSEI
            For BAC library availability, please contact Masa-Toshi Yamamoto
            (yamamoto@kit.jp).
            Submitted (30-11-2004) by Masahira Hattori,
            RIKEN, Genomic Sciences Center (GSC);
            1-7-22 Saitama-shi, Tsukuba-shi, Ibaraki, Japan, 305-3858, Japan
            (E-mail:hattori@gs.c.riken.jp, Tel:81-45-503-9111,
            Fax:81-45-503-9170)
            This work was done in collaboration with Yamamoto, M.-T. Drosophila
            Genetic Resource Center
            Suga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
            Tel: 81-75-873-2660 FAX: 81-75-861-0881
            PRIMERS
            Sequencing : F
LIBRARY
Vector      : pKS150
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .623
   /organism="Drosophila sechellia"
   /mol_type="genomic DNA"
   /db_xref="taxon:7238"
   /clone="DSEI-004A05.F.fa"
   /clone_lib="DSEI Drosophila BAC library"
FEATURES
SOURCE
Query Match      4.5%; Score 62.2; DB 14; Length 623;
Best Local Similarity 51.0%; Pred. No. 3.7e-06;
Matches 176; Conservative 0; Mismatches 163; Indels 6; Gaps 1;
ORIGIN
D 703 GCATTTGCTGGCTAAATTGGCATGCTTGTGATAGATGCGGCTTTGGCAATTTTCAA 762
Y 703 GCATTTGCTGGCTAAATTGGCATGCTTGTGATAGATGCGGCTTTGGCAATTTTCAA 762
D 285 GAATATATCAGCAATATCTCATGTGCATCTATTCAGGCAACCTCACAACGCTTAT 344
Y 285 GAATATATCAGCAATATCTCATGTGCATCTATTCAGGCAACCTCACAACGCTTAT 344
D 763 AATACGCAATTTTGTGCAATAATCCACATCCATCCGCGCATCAAGATATCAGC 822
Y 763 AATACGCAATTTTGTGCAATAATCCACATCCATCCGCGCATCAAGATATCAGC 822
D 345 CAGTTTAAATTAATGTTGATAGTGTCTCATCTTCATCCATCTATCTATGATCTGG 404
Y 345 CAGTTTAAATTAATGTTGATAGTGTCTCATCTTCATCCATCTATCTATGATCTGG 404
D 823 GAGAGAGTTTAAATTTGTTGCCCTTGAGTGCCTCAATTTGATATGCTCAAGCTGAA 882
Y 823 GAGAGAGTTTAAATTTGTTGCCCTTGAGTGCCTCAATTTGATATGCTCAAGCTGAA 882
D 405 TGTAAATTTGAGCATGTGTAATGTCCAGTAAATTTGATTCATCAATCTCAATTTGAA 464
Y 405 TGTAAATTTGAGCATGTGTAATGTCCAGTAAATTTGATTCATCAATCTCAATTTGAA 464
D 883 ACTTGAAGCTCAAGACCAACCAATTCATATTTTGATCGTAAATTTGAAGTCAGG 942
Y 883 ACTTGAAGCTCAAGACCAACCAATTCATATTTTGATCGTAAATTTGAAGTCAGG 942
D 465 AGAGAGAGTTCGATTCATATAATTTTGCATCTC-----TTTCTGATCAAAACAGG 518
Y 465 AGAGAGAGTTCGATTCATATAATTTTGCATCTC-----TTTCTGATCAAAACAGG 518
D 943 CTTGCGCAATATTTACCGCGGACCAACAATGATGCTGATGCTTTGCAATCTGCGCTAC 1002
Y 943 CTTGCGCAATATTTACCGCGGACCAACAATGATGCTGATGCTTTGCAATCTGCGCTAC 1002
D 519 AACGCTTAATTTCCACCAATAGCTATTTTTCCTGCAAGATTTTAATATGTCCTAT 578
Y 519 AACGCTTAATTTCCACCAATAGCTATTTTTCCTGCAAGATTTTAATATGTCCTAT 578

```

```

ORIGIN
D 1003 TAAATGCTAACAAGCTCTTTGGCATTTGAGCCGTGATGCGCAT 1047
Y 1003 TAAATGCTAACAAGCTCTTTGGCATTTGAGCCGTGATGCGCAT 1047
D 579 TAGATGATGATGCTGATTTACCATTTGTTCTCTGACCGCTAT 623
Y 579 TAGATGATGATGCTGATTTACCATTTGTTCTCTGACCGCTAT 623
RESULT 10
DUT69542
LOCUS
DEFINITION  928 bp DNA linear GSS 27-JAN-2006
            ARK61505.b2 HF4000_12-21-03 uncultured marine microorganism
            HF4000_12-21-03 genomic clone HF4000_13841003B18, genomic survey
            sequence.
ACCESSION  DUT69542
VERSION    DUT69542
KEYWORDS   GSS.
SOURCE     DUT69542.1  GI:85783725
ORGANISM   uncultured marine microorganism HF4000_12-21-03
            unclassified sequences; environmental samples.
REFERENCE  1 (bases 1 to 928)
AUTHORS   Delong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
TITLE      Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior
JOURNAL    Science (2006) In press
COMMENT    Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
            Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
            and Edward Delong
            US DOE Joint Genome Institute
            2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
            Tel: 617-253-5271
            Fax: 617-253-2679
            Email: PMRichardson@lbl.gov; delong@mit.edu
            North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
            DNA library prepared from marine picoplankton in the less than 1.6
            um, greater than 0.22 um fraction. Sample Date: 12/21/2003
            Coordinates: 22.45 N, 158 W Depth 4000 m Temperature: 1.46 C
            Salinity: 34.69 Oxygen: 147.8 umol/kg
            Class: fosmid ends.
FEATURES
SOURCE
1. .928
   /organism="uncultured marine microorganism
   HF4000_12-21-03"
   /mol_type="genomic DNA"
   /db_xref="taxon:361148"
   /clone="HF4000_13841003B18"
   /cell_type="marine picoplankton, less than 1.8 um, greater
   than 0.22 um fraction"
   /clone_lib="HF4000_12-21-03"
   /notes="Vector: pCC1FOS; North Pacific Subtropical Gyre
   (Hawaii) picoplankton genomic fosmid DNA library prepared
   from marine picoplankton in the less than 1.6 um, greater
   than 0.22 um fraction. picoplankton collected at 4000 m
   depth on 10/7//2002, Coordinates: 22.45 N, 158 W. Sample
   Date: 12/21/2003 Coordinates: 22.45 N, 158 W Depth 4000 m
   Temperature: 1.46 C Salinity: 34.69 Oxygen: 147.8 umol/kg"
ORIGIN
Query Match      4.3%; Score 59.6; DB 14; Length 928;
Best Local Similarity 50.3%; Pred. No. 2.3e-05;
Matches 146; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
ORIGIN
D 758 TCAAAATAGCAATTTTGTGCAATAATCCACCATGTCATCGTCAAGATGA 817
Y 758 TCAAAATAGCAATTTTGTGCAATAATCCACCATGTCATCGTCAAGATGA 817
D 143 TAAAGATCATCTTTCTGGGGGAGATATCTGCAATGCGCGTAAACGCTCTGATG 202
Y 143 TAAAGATCATCTTTCTGGGGGAGATATCTGCAATGCGCGTAAACGCTCTGATG 202
D 818 TCAAGCGAGAGATTTAAATTTGTTGCCCTTGAGCTGCCAAATTTGATATGCTCAAGC 877
Y 818 TCAAGCGAGAGATTTAAATTTGTTGCCCTTGAGCTGCCAAATTTGATATGCTCAAGC 877
D 203 TCGGGGCTGATGTTCAACACACAGGAGCTCGAAGTGTATCGAACGGGTCAATTCGAGC 262
Y 203 TCGGGGCTGATGTTCAACACACAGGAGCTCGAAGTGTATCGAACGGGTCAATTCGAGC 262
D 878 TGAATACTTGAACAGCTAAGACCAACAATTCATATTTTGATCGTAAATATTCAGT 937
Y 878 TGAATACTTGAACAGCTAAGACCAACAATTCATATTTTGATCGTAAATATTCAGT 937

```

Db 263 TGAAGACGACATCTCCAGACATAGATGCGCTTGTTCACAGGGATCCATGTCAGC 322  
Qy 938 GAGGGCTGGCAATATTACCGCGACACAAACATATGCTGTGATGTTTGCATCTCG 997  
Db 323 GCGGGATACCCAGATGCGCGCATTTTCGCTCGCGCGACACCTGCATGATGTG 382  
Qy 998 CCTACTATGCTGTAACAGTCTTGTTCGATTTGAGCTGTGATGGCAT 1047  
Db 383 CCGATCAACAGACGTCGTGATTTGCCGTGTGCGGTAATGCCGAT 432

RESULT 11  
CL691264/c 628 bp DNA linear GSS 10-JUL-2004  
LOCUS PR10156a.A05.2 - PR10156a.BR (628) Mixed stage fosmid library of P.  
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.  
ACCESSION CL691264 GI:50213172  
VERSION CL691264  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasterida; Pristionchus.  
REFERENCE 1 (bases 1 to 628)  
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.  
AppaDB: An Acids database for the nematode satellite organism  
Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
14681447  
JOURNAL Contact: Sommer RJ  
PUBMED Evolutionary Biology  
COMMENT Max-Planck-Institute for Developmental Biology  
Semannstr. 37-39, Tuebingen D-72076, Germany  
Tel.: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end  
sequenced at Vancouver, Canada.  
Seq primer: T7  
Class: fosmid ends.  
Location/Qualifiers  
1. 628  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage fosmid library of P. pacificus  
var. California"  
/note="Vector: pEpfos-5 Fosmid vector"

ORIGIN  
Query Match 4.1%; Score 56.8; DB 13; Length 628;  
Best Local Similarity 48.0%; Pred. No. 0.00014;  
Matches 246; Conservative 0; Mismatches 242; Indels 24; Gaps 2;

Qy 1 TAAAGCTGAACCATGTAACAACTATACCTGCTCATATTAATTTTAATTATC 60  
Db 494 TAACTCTTGCGCAGACGGGCAACTATGCTGTTGTTCAAGTTCTTGAATCATC 435

Qy 61 CATGCTAGCACAAGCAGGTGACAAAGACAGCCCTAGCGGTGTTGATGCTGCTCA 120  
Db 434 AAGGCTTGACACAGGCTGGGAGAGACAAACAT-----ATCGCCGGCTG 390  
Qy 121 TACTGCCGATCAAGGTATACCACTGATATTTAGATGCCAATGTGCATCACTTT 180  
Db 389 AACACGGGAGAGCGCAAGCGCATCGCTGTTCCATGTTTGGTTGTTCTGCACTTC 330  
Qy 181 GAGTAATCTTGCTCATTTATCGGGGATTTGACCGATTAATAACACATATCCACCA 240  
Db 329 CGGGCGTAGCGCGCCAGCTGCGCGGCTGACGACCAAAACAATAGAGTAAGTTATC 270  
Qy 241 CTCGATTAACATATGCTGACAGCTCAATAATCTTGACCTTTACCAAGCCCAAAAT 300

Db 269 GCCATTAGGTAAAGCGGCCAGTGGGCTAAAGTCCGGCATTTACATCGCACCCACAA 210  
Qy 301 CAATGCCAGTGATGATATGCCATTAACCGTCCAAAGCCATCAATTTCTGATGTGA 360  
Db 209 CAATGACAGTGTGCGGT-----CTAGTCCAGGCCATTTACAGCGCCCTTCCTGACT 159  
Qy 361 GCCGATTTGTGACCTTTGATCATTAATAATATGCTTGCATATCATCATATATA 420  
Db 158 GCCGACCTTGATGCTTTTCATATGTTATCCAAACGATACCGCTTATGCTCCAGACAC 99

Qy 421 TTCAACGCGATGCGGTAGCCCTTTGAAAGCTTAAAGGTATCAAGCATGATCATGAG 480  
Db 98 TTCAAGCGATGCGGACGACGACGATGATGTTAAGCTTTACAGCTGTGCGACGCG 39  
Qy 481 AAGATTGCGAGCTGTGCTTAAAGCAGTGCAG 512  
Db 38 TAACTTCGACGACATCTGCCAGCCGCGCGCG 7

RESULT 12  
DR032465/c 474 bp mRNA linear EST 31-MAY-2005  
LOCUS bda020270D15.abl.061 Antrodia cinnamomea cDNA library Antrodia  
DEFINITION cinnamomea cDNA, mRNA sequence.  
ACCESSION DR032465 GI:66834715  
VERSION DR032465.1 GI:66834715  
KEYWORDS EST.  
SOURCE Antrodia cinnamomea  
ORGANISM Antrodia cinnamomea  
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
Aphyllophorales; Antrodia.  
REFERENCE 1 (bases 1 to 474)  
Tzean,S.S., Hwang,S.T., Song,H.C., Liou,R.F., Yehk,H.H. and  
Shiao,M.S.  
Analysis of expressed sequence tags from Antrodia cinnamomea, a  
polypora causing brown rot of Cinnamomum kanehirai  
unpublished (2005)  
JOURNAL Contact: Tzean SS  
COMMENT Department of Plant Pathology and Microbiology  
National Taiwan University  
NO.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan  
Tel: 886 2 3366 4595  
Fax: 886 2 2362 0639  
Email: setentn.edu.tw  
PCR Primers  
FORWARD: 5' GTAAGACGAGCCAGT 3'  
REVERSE: 5' CACACAGAAACGATATGACCAT 3'.  
BACKWARD: 5' CACACAGAAACGATATGACCAT 3'.  
Location/Qualifiers  
1. 474  
/organism="Antrodia cinnamomea"  
/mol\_type="mRNA"  
/strain="SMV 1"  
/db\_xref="taxon:279009"  
/sex="Heterokaryon"  
/cell\_type="Mycelium"  
/lab\_host="Cinnamomum kanehirai"  
/clone\_lib="Antrodia cinnamomea cDNA library"  
/note="Vector: ZAP Express Vector"

ORIGIN  
Query Match 4.0%; Score 55; DB 9; Length 474;  
Best Local Similarity 54.1%; Pred. No. 0.00042;  
Matches 112; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1171 AATGCGGGGCTGATGATGATGATCGGTGAGTTAAGAGTTCACTGTCAAGCTGCC 1230  
Db 453 AATACCGGAGCTGGCGACATCATGCTGCCGATCAGCATTCATTCATTCAGACTGCC 394  
Qy 1231 AAAATGACATGAAGCGCTGTGGCGAGTTTATTCGCAAGGTAGATGGGATTTACATC 1290  
Db 393 CGTGTGGCGTTCTTACGCGCTTGGGTAAATTTATTCAGAGCGGCGGTGTCTATACGCTATC 334

| QY                        | 1251  | GATGATGCTAACTTGTGTGACCGTGATTTGACCAAAAAATTCCAGCTGCACGTCCAGAGCT | 1350            |
|---------------------------|---|---|-----------------|
| Db                        | 333   | CATAACGGCGCGCGTCACACCGCGAGCGAGGAAAAAGTCCACGACGAAAAAGCCGGTGA   | 274             |
| QY                        | 1351  | GGCCAAACCGACGACGCAATATTTTTT                                   | 1377            |
| Db                        | 273   | GGCCAGCGCGATATATGACGACATTTTTT                                 | 247             |
| RESULT 13                 |   |   |                 |
| DU732353/c                |   |   |                 |
| LOCUS                     |   |   |                 |
| DEFINITION                | APK1690 b2 HF70_10-07-02 uncultured marine microorganism                | linear  | SSS 27-JAN-2006 |
| ACCESSION                 | HF70_10-07-02 genomic clone HF0070_025B08, genomic survey sequence.     |   |                 |
| VERSION                   | DU732353  |   |                 |
| KEYWORDS                  | DU732353.1 GI:85742187  |   |                 |
| SOURCE                    | SSS.  |   |                 |
| ORGANISM                  | uncultured marine microorganism HF70_10-07-02                           |   |                 |
| REFERENCE                 | uncultured marine microorganism HF70_10-07-02                           |   |                 |
| AUTHORS                   | unclassified sequences; environmental samples.                          |   |                 |
|                           | 1 (bases 1 to 1024)   |   |                 |
|                           | Delong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,        |   |                 |
|                           | Frigaard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W. |   |                 |
|                           | and Karl, D.M.  |   |                 |
| TITLE                     | Comparative genomics reveals ecological trends in stratified            |   |                 |
| JOURNAL                   | microbial communities in the ocean's interior                           |   |                 |
| COMMENT                   | Science (2006) In press   |   |                 |
|                           | Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,         |   |                 |
|                           | Kerrie Barry, Tijana Glavinadze, David Bruce, Paul Richardson           |   |                 |
|                           | US DOE Joint Genome Institute   |   |                 |
|                           | US DOE Joint Genome Institute   |   |                 |
|                           | 2800 Mitchell Drive Bldg B100, Walnut Creek, CA 94598-1698, USA         |   |                 |
|                           | Tel: 617-253-5271   |   |                 |
|                           | Fax: 617-253-2679   |   |                 |
|                           | Email: PMRichardson@lbl.gov; delong@mit.edu                             |   |                 |
|                           | Sample Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth 70 m           |   |                 |
|                           | Temperature: 24.93 C Salinity: 35.21 psu Oxygen: 217.4 umol/Kg          |   |                 |
|                           | Class: fosmid ends.   |   |                 |
| FEATURES                  |   |   |                 |
| source                    | Location/Qualifiers   |   |                 |
|                           | 1..1024   |   |                 |
|                           | /organism="uncultured marine microorganism HF70_10-07-02"               |   |                 |
|                           | /mol_type="genomic DNA"   |   |                 |
|                           | /db_xref="taxon:361144"   |   |                 |
|                           | /clone="HF0070_025B08"  |   |                 |
|                           | /cell_type="marine picoplankton, less than 1.8 um, greater              |   |                 |
|                           | than 0.22 um fraction"  |   |                 |
|                           | /clone_1lb="HF70_10-07-02"  |   |                 |
|                           | /note="Vector: pCICFOS; North Pacific Subtropical Gyre                  |   |                 |
|                           | (Hawaii) picoplankton genomic fosmid DNA library prepared               |   |                 |
|                           | from marine picoplankton in the less than 1.6 um, greater               |   |                 |
|                           | than 0.22 um fraction. Picoplankton collected at 70 m                   |   |                 |
|                           | depth on 10/7/2002. Coordinates: 22.45 N, 158 W Sample                  |   |                 |
|                           | Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth 70 m                  |   |                 |
|                           | Temperature: 24.93 C Salinity: 35.21 psu Oxygen: 217.4                  |   |                 |
|                           | umol/Kg"  |   |                 |
| ORIGIN                    |   |   |                 |
| Query Match               | 3.9%  | Score 54.2; DB 14; Length 1024;                               |                 |
| Best Local Similarity     | 51.4%   | Pred. No. 0.00088;  |                 |
| Matches 125; Conservative | 0; Mismatches 118; Indels   | 0; Gaps   | 0;              |
| QY                        | 333   | CAAGCCATCAATTTGCTGCAATGTTGAGCCGATGTTTGATCCTTTTGATATCAATTAAT   | 392             |
| Db                        | 557   | CCAGGCGCTCCAGCGCGCGCGCGTGCACCGATATTTGCTGCTTGTGCTGATGATGA      | 498             |
| QY                        | 393   | ATGCTTGTATCATATATCATATATTTACAGCGATGCGTGACCTTTGAAAAGTCT        | 452             |
| Db                        | 497   | ACGCGACGCGCTTTGACTTGAATTAAGCTGCGACGAGGAGGCGACACTCGCATCTCT     | 438             |
| QY                        | 453   | TAAAGGTATCAAGCATGAGCATCATGAGGAATGCGAGCTGTGCTTAAGCAAGTGA       | 512             |
| Db                        | 437   | TGTGCTCCCTGATGATGCTCCCTGAGGCGCATTTTCAACGCGCTGCCCATGATCAAGCGCG | 378             |

[illegible]

Db 99 ACGGACAGCGGACCGATGGCCAGTGCCTCGGAGACGGCTTG 53

Job time : 6248 secs

RESULT 15  
CZ545153 838 bp DNA linear GSS 13-MAY-2005  
LOCUS SRAA-aad60c01.g1 Strongyloides ratii whole genome shotgun library  
DEFINITION (SRAAGSS 004) Strongyloides ratii genomic, genomic survey sequence.  
ACCESSION CZ545153  
VERSION CZ545153.1 GI:64677289  
KEYWORDS GSS.  
SOURCE Strongyloides ratii  
ORGANISM Strongyloides ratii  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Panagrolaimidae; Strongyloidea; Strongyloides.

REFERENCE  
AUTHORS 1 (bases 1 to 838)  
Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Page,D.,  
Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,  
Clifton,S.W. and Wilson,R.  
TITLE Genome Survey sequences from the rat parasitic nematode  
Strongyloides ratii  
JOURNAL Unpublished (2005)  
COMMENT Contact: Mitreva M  
Washington University in St. Louis  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: nematode@watson.wustl.edu  
Genomic DNA was provided by Fiona Thompson  
(F.Thompson@bristol.ac.uk) and Mark Viney  
(Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,  
UK.

Class: shotgun.

Location/Qualifiers  
1..838  
/organism="Strongyloides ratii"  
/mol\_type="genomic DNA"  
/strain="Isotemale line ED321 heterogonic"  
/db\_xref="taxon:34506"  
/dev\_stage="infective larval stage (L3)"  
/lab\_host="GS10"  
/clone\_lib="Strongyloides ratii whole genome shotgun  
library (SRAAGSS 004)"  
/note="Vector: POTW13; Site 1: BstXI; Site 2: BstXI;  
Strongyloides ratii genomic DNA was randomly sheared,  
end-repaired and size fractioned to enrich for 2-4 kb  
fragments. Genomic DNA was provided by Fiona Thompson  
(F.Thompson@bristol.ac.uk) and Mark Viney  
(Mark.Viney@bristol.ac.uk) at the University of Bristol,  
Bristol, UK. Sequencing by Washington University Genome  
Sequencing Center, St. Louis, MO."

ORIGIN

Query Match 3.6%; Score 50.2; DB 13; Length 838;  
Best Local Similarity 62.2%; Pred. No. 0.012;  
Matches 79; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 946 GCCAATATTACCGCCGACCAACAATCGCTGATGTTTGGCTTCGCTACTAA 1005  
DB 76 GCCAATATTTCGCCCTTAACCAATTCATCTCTTTTTCAGAGATGATAGATTAA 135  
OY 1006 TGTGTAACAGTCTTTTGGCATTGAGCCTGTGATGCGAGATGATGTTGTTGGC 1065  
DB 136 AGACGTTGTGTTGTTTTCGTTGCTTCTGTGATGCAATGATTTTCGATCGGTA 195  
OY 1066 TTGAGTT 1072  
DB 196 TTCAGAT 202

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2006, 09:25:31 ; Search time 255 Seconds  
(without alignments)  
10126.001 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380  
Sequence: 1 taagcgtagacatctta.....acgagcgcatatttttga 1380

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2268460

Minimum DB seq length: 22

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/1\_COMB.seq: \*  
2: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/5\_COMB.seq: \*  
3: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/6\_COMB.seq: \*  
4: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/7\_COMB.seq: \*  
5: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/7\_COMB.seq: \*  
6: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/H\_COMB.seq: \*  
7: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/PC/US\_COMB.seq: \*  
8: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/PC\_COMB.seq: \*  
9: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/RE\_COMB.seq: \*  
10: /EMC\_Celerra\_SIDS3/Pcdata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length  | ID | Description                             |
|------------|--------|-------------|---------|----|---|
| 1          | 1380   | 100.0       | 96109   | 3  | US-09-596-002-35 Sequence 35, Appl      |
| 2          | 1370.4 | 99.3        | 1410    | 3  | US-09-540-236-673 Sequence 673, Appl    |
| 3          | 1226.2 | 16.4        | 1416    | 3  | US-09-328-352-1282 Sequence 1282, Appl  |
| 4          | 160.4  | 11.6        | 1830121 | 3  | US-09-557-884-1 Sequence 1, Appl        |
| 5          | 160.4  | 11.6        | 1830121 | 3  | US-09-643-990A-1 Sequence 1, Appl       |
| 6          | 160.4  | 11.6        | 1830121 | 3  | US-10-158-865-1 Sequence 1, Appl        |
| 7          | 125.2  | 9.1         | 1373    | 3  | US-09-543-681A-2830 Sequence 2830, Appl |
| 8          | 117    | 8.5         | 1371    | 3  | US-09-252-991A-7702 Sequence 7702, Appl |
| 9          | 117    | 8.5         | 1401    | 3  | US-09-252-991A-7861 Sequence 7861, Appl |
| 10         | 117    | 7.3         | 1450    | 3  | US-09-701-829-1 Sequence 1, Appl        |
| 11         | 100.6  | 7.3         | 1416    | 3  | US-09-489-039A-5247 Sequence 5247, Appl |
| 12         | 84.8   | 6.1         | 640681  | 3  | US-09-790-988-1 Sequence 1, Appl        |
| 13         | 83     | 6.0         | 1368    | 3  | US-09-107-532A-1574 Sequence 1574, Appl |
| 14         | 69.8   | 5.1         | 1567    | 3  | US-09-252-991A-7787 Sequence 7787, Appl |
| 15         | 67     | 4.9         | 2402    | 3  | US-09-221-017B-881 Sequence 881, Appl   |
| 16         | 61.4   | 4.4         | 1372    | 3  | US-09-530-836-1 Sequence 1, Appl        |
| 17         | 58.6   | 4.2         | 1242    | 3  | US-09-147-928-1 Sequence 1, Appl        |
| 18         | 53.6   | 3.9         | 2423    | 3  | US-08-714-918-86 Sequence 86, Appl      |
| 19         | 53.6   | 3.9         | 2423    | 3  | US-09-265-315-86 Sequence 86, Appl      |
| 20         | 53.6   | 3.9         | 2423    | 3  | US-09-265-315-86 Sequence 86, Appl      |
| 21         | 53.6   | 3.9         | 2423    | 3  | US-09-266-417-86 Sequence 86, Appl      |
| 22         | 53.6   | 3.9         | 2423    | 3  | US-09-528-709-86 Sequence 86, Appl      |
| 23         | 53.6   | 3.9         | 2423    | 3  | US-09-527-745-86 Sequence 86, Appl      |

|      |      |     |       |   |                     |                     |
|------|------|-----|-------|---|---------------------|---------------------|
| C 24 | 51.8 | 3.8 | 1352  | 5 | US-09-974-300-1685  | Sequence 1685, Appl |
| C 25 | 49.6 | 3.6 | 536   | 3 | US-09-147-928-3     | Sequence 3, Appl    |
| C 26 | 49.4 | 3.6 | 699   | 3 | US-09-134-000C-741  | Sequence 741, Appl  |
| C 27 | 48.6 | 3.5 | 1520  | 3 | US-08-956-171E-80   | Sequence 80, Appl   |
| C 28 | 48.6 | 3.5 | 1320  | 3 | US-08-781-986A-80   | Sequence 80, Appl   |
| C 29 | 48.2 | 3.5 | 1188  | 3 | US-09-710-279-655   | Sequence 655, Appl  |
| C 30 | 48.2 | 3.5 | 3334  | 3 | US-09-710-279-3574  | Sequence 3574, Appl |
| C 31 | 47   | 3.4 | 1374  | 3 | US-09-134-001C-1001 | Sequence 1001, Appl |
| C 32 | 45.6 | 3.3 | 405   | 5 | US-09-974-300-6173  | Sequence 6173, Appl |
| C 33 | 45.6 | 3.3 | 717   | 3 | US-08-936-165A-78   | Sequence 78, Appl   |
| C 34 | 42.6 | 3.1 | 1368  | 3 | US-09-902-540-9337  | Sequence 9337, Appl |
| C 35 | 42.6 | 3.1 | 10486 | 3 | US-09-902-540-997   | Sequence 997, Appl  |
| C 36 | 42   | 3.0 | 1407  | 3 | US-09-530-838-1     | Sequence 1, Appl    |
| C 37 | 41.8 | 3.0 | 6693  | 3 | US-08-961-527-195   | Sequence 195, Appl  |
| C 38 | 41   | 3.0 | 4010  | 3 | US-09-710-279-3540  | Sequence 3540, Appl |
| C 39 | 40.8 | 3.0 | 1141  | 3 | US-09-806-708B-22   | Sequence 22, Appl   |
| C 40 | 40.8 | 2.9 | 1275  | 3 | US-10-113-113-3     | Sequence 3, Appl    |
| C 41 | 40.2 | 2.9 | 1350  | 2 | US-08-665-435A-1    | Sequence 1, Appl    |
| C 42 | 40.2 | 2.9 | 1350  | 2 | US-08-665-435A-3    | Sequence 3, Appl    |
| C 43 | 40.2 | 2.9 | 1350  | 2 | US-08-843-309-3     | Sequence 3, Appl    |
| C 44 | 40.2 | 2.9 | 1353  | 2 | US-08-843-309-1     | Sequence 1, Appl    |
| C 45 | 40.2 | 2.9 | 1353  | 3 | US-09-583-110-197   | Sequence 197, Appl  |

#### ALIGNMENTS

RESULT 1  
US-09-596-002-35  
Sequence 35, Application US/09596002  
Patent No. 6632636  
GENERAL INFORMATION:  
APPLICANT: Lagace, Robert, E.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
FILE REFERENCE: PM-0008-4 US  
CURRENT APPLICATION NUMBER: US/09/596, 002  
PRIOR FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: 60/140,121  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PERL Program  
SEQ ID NO 35  
LENGTH: 96109  
TYPE: DNA  
ORGANISM: M. catarrhalis  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte template ID No. 6632636 35  
PUBLICATION INFORMATION:  
US-09-596-002-35

Query Match 100.0%; Score 1380; DB 3; Length 96109;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |       |   |       |
|----|-------|---|-------|
| QY | 1     | TAAGCGTAAACATGTTTAAACAACTTACCTGCTTATTAATTTTAAATGATC     | 60    |
| DB | 11357 | TAAGCGTAAACATGTTTAAACAACTTACCTGCTTATTAATTTTAAATGATC     | 11416 |
| QY | 61    | CATGCGAGCAAGCAAGGAGCAAAAGCAACCGCTAGGCGTCTTGACCTGGCTGCAA | 120   |
| DB | 11417 | CATGCGAGCAAGCAAGGAGCAAAAGCAACCGCTAGGCGTCTTGACCTGGCTGCAA | 11476 |
| QY | 121   | TACTGCCGATCAAGTGTACCAACCTGATATTTTAAATGCAATGTCATCACTTT   | 180   |
| DB | 11477 | TACTGCCGATCAAGTGTACCAACCTGATATTTTAAATGCAATGTCATCACTTT   | 11536 |
| QY | 181   | GAGTAATCTTGGCTCAATTATGGGGCATCTTGACCGATTAATACACTATCCACCA | 240   |
| DB | 11537 | GAGTAATCTTGGCTCAATTATGGGGCATCTTGACCGATTAATACACTATCCACCA | 11596 |

QY 241 CTGCATACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAAT 300  
Db 11597 CTGCATACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAAT 11656  
QY 301 CAATGCCAGTATGATATGCTCAATPAAACCGTGCCAAAGCCCATCAATGCTGCAATGTTGA 360  
Db 11657 CATGCGCAGTATGATATGCTCAATPAAACCGTGCCAAAGCCCATCAATGCTGCAATGTTGA 11716  
QY 361 GCGGATGTTGTAACCTTTGTAATCATTAATAATATGCTGCTATCAATATCATCAATATA 420  
Db 11717 GCGGATGTTGTAACCTTTGTAATCATTAATAATATGCTGCTATCAATATCATCAATATA 11776  
QY 421 TTCACAGCGATGCGGTAGCCCTTTGAAGGCTTAAAGGCTTCAAGCATGCGATCAAGG 480  
Db 11777 TTCACAGCGATGCGGTAGCCCTTTGAAGGCTTAAAGGCTTCAAGCATGCGATCAAGG 11836  
QY 481 AAGATTGGCAGCTGTGCTTAAAGCAGTCAAGTAAAGGCTTGAAGTAAAGTAAAGTAAAG 540  
Db 11837 AAGATTGGCAGCTGTGCTTAAAGCAGTCAAGTAAAGGCTTGAAGTAAAGTAAAGTAAAG 11896  
QY 541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 11897 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11956  
QY 601 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
Db 11957 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12016  
QY 661 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 12017 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12076  
QY 721 GGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Db 12077 GGCATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12136  
QY 781 CAATTAATCACCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db 12137 CAATTAATCACCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12196  
QY 841 TGGCCCTTGAAGCTGCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 12197 TGGCCCTTGAAGCTGCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 12256  
QY 901 CACCAATCATATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 12257 CACCAATCATATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12316  
QY 961 GACACCAATCATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 12317 GACACCAATCATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12376  
QY 1021 TTTGGCATTTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Db 12377 TTTGGCATTTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12436  
QY 1081 TGGCTTGAAGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140  
Db 12437 TGGCTTGAAGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12496  
QY 1141 GCGGCGACAAATCTCGGCGCTTCTTGCGTTAAATGCGGCGCTGATGATGATGATGATG 1200  
Db 12497 GCGGCGACAAATCTCGGCGCTTCTTGCGTTAAATGCGGCGCTGATGATGATGATGATGAT 12556  
QY 1201 TGAAGCTTAAAGTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
Db 12557 TGAAGCTTAAAGTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12616  
QY 1261 ATTCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
Db 12617 ATTCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 12676  
QY 1321 CAAAAAATTCACAGCTGACAGTCCAGCTGCCCAACCGACGACGCAATATTTTGTGA 1380

Db 12677 CAAAAAATTCACAGCTGACAGTCCAGCTGCCCAACCGACGACGCAATATTTTGTGA 12736  
RESULT 2  
US-09-540-236-673c  
; Sequence 673, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breco et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRI  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 673  
; LENGTH: 1410  
; TYPE: DNA  
; ORGANISM: M. catarrhalis  
US-09-540-236-673  
Query Match 99.3%; Score 1370.4; DB 3; Length 1410;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1374; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 TAAAGCGTAAACCATGTTAAACAACTTATCAGCTGCTATTAATTTTAAATGATC 60  
Db 1392 TAAAGCGTAAACCATGTTAAACAACTTATCAGCTGCTATTAATTTTAAATGATC 1333  
QY 61 CATGCTAGCAACAAGCAGTGAACAAAGCAAGCCCTAGGCGTCTTGAAGTGGCCAA 120  
Db 1332 CATGCTAGCAACAAGCAGGCGAACAAGGCAAGCCCTAGGCGTCTTGAAGTGGCCAA 1273  
QY 121 TACTGCCGATCAAGTGTACCAACCTGATATTTTAAATGCCAAATGTCATCCTTT 180  
Db 1272 TACTGCCGATCAAGTGTACCAACCTGATATTTTAAATGCCAAATGTCATCCTTT 1213  
QY 181 GAGTAATCTGTGCTCAATATGCGGGCATCTTGACCGATTAATAACACTATCCACCA 240  
Db 1212 GAGTAATCTGTGCTCAATATGCGGGCATCTTGACCGATTAATAACACTATCCACCA 1153  
QY 241 CTGCATACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAAT 300  
Db 1152 CTGCATACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAAT 1093  
QY 301 CAATGCCAGTATGATATGCTCAATPAAACCGTGCCAAAGCCCATCAATGCTGCAATGTTGA 360  
Db 1092 CAATGCCAGTATGATATGCTCAATPAAACCGTGCCAAAGCCCATCAATGCTGCAATGTTGA 1033  
QY 361 GCGGATGTTGTAACCTTTGTAATCATTAATAATATGCTGCTATCAATATCATCAATATA 420  
Db 1032 GCGGATGTTGTAACCTTTGTAATCATTAATAATATGCTGCTATCAATATCATCAATATA 973  
QY 421 TTCACAGCGATGCGGTAGCCCTTTGAAGGCTTAAAGGCTTCAAGCATGCGATCAAGG 480  
Db 972 TTCACAGCGATGCGGTAGCCCTTTGAAGGCTTAAAGGCTTCAAGCATGCGATCAAGG 913  
QY 481 AAGATTGGCAGCTGTGCTTAAAGCAGTCAAGTAAAGGCTTGAAGTAAAGTAAAGTAAAG 540  
Db 912 AAGATTGGCAGCTGTGCTTAAAGCAGTCAAGTAAAGGCTTGAAGTAAAGTAAAGTAAAG 853  
QY 541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 852 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793  
QY 601 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660  
Db 792 ATCACCATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 733  
QY 661 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 732 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 673





Db 329 GCAATAGCTGCTGAAATTTCCGGTAATTTGGTGCAAGGCCCTGGCTTAAATATCTCT 270  
Qy 1196 TCGGCTGAGCTTAAGATTCATCTGTCCAGCTGCCAAATGACATGACGCCCTGTGGGC 1255  
Db 269 TCTGCTTGTATTAATTAATTTCTTGATCAAGCTGCAAAACTGGTTTAAACCAACGCGGA 210  
Qy 1256 AGTTTATCTGCCAAGTAGGATTTGGATTAACATGATGATGCTAACTTTGTGACCGTGA 1315  
Db 209 ATCTGATCTGTGCGGAGGTGTGGGCGGGAATCCGTTACAGCAACTTGTGTAGCCCTGT 150  
Qy 1316 TTGACCAAAAATTTCAAGCTGACAGTCCAGAGCTGCCAAACCGAGAGAGGATAT 1372  
Db 149 TCATGACGAAATTTACAGCAAAACCTGTAATTTCCCAAGCCTGTACAACTTTT 93

RESULT 4  
US-09-557-884-1/C  
Sequence 1, Application US/09557884  
Patent No. 6506581  
GENERAL INFORMATION:  
APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557, 884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Query Match 11.6%; Score 160.4; DB 3; Length 1830121;  
Best Local Similarity 49.6%; Pred. No. 9.6e-37;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy 1 TAAAGCTGAGACATGTTAACTATACCTGCTCATTTATTTTAAATTTTAAATGATC 60  
Db 1204439 TAAACATTTAGCTTAATGCTTAATTTCTTGCGCGGCTTTTCAAAACAAACGATC 1204380  
Qy 61 CATGCTAGCAACAAGCAGGTGACAAAGACAGCCCTAGCGGTGCTTTGACTGCTGCCAA 120  
Db 1204379 GAGACTTGCAACAAGCAGGTGACAAATACCA-----TATCTCGGCTTTGCA 1204332  
Qy 121 TACTGCCGATCAAGTATACCACTGATATTTTATGATGCCAAATGTGATCACCCTT 180

Db 1204331 TGTGGGCGGTAATAATTTCTATCGCTTGTTCATTTGATGAAACA-----ATACTTTG 1204278  
Qy 181 GAGTAATCTGTGCTAATTAATGGGGCATCTGACCGATTAATATACACTATCCACCA 240  
Db 1204277 CGATGAAAATTTTGGACAGCGGCACACTTCGACCAAAACATTAACATTAATGTGTGG 1204218  
Qy 241 CTGATTAACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAAT 300  
Db 1204217 TTGATTAATTAATTCAGCTAATTTCTGAAAATTCAGCCCTTTCCGCTCCGCTAGCA 1204158  
Qy 301 CAATGCCAGTATGATATGCAATAACCGTGCCAAAGCCCATCAATTTCTGCAATGTTGA 360  
Db 1204157 CAATGCAATTTACCCCTCAATATA-----AGCCAGGCAATGACAGCAACTGTACT 1204107  
Qy 361 GCGGATGTTGTAACCTTTGTAATCATTAATAATATGCTTGATCAATATCATATATA 420  
Db 1204106 CCCACATTTGTGTCTTATAGATCATTAATCCAGCAATGCCATTTAGCTTGATGACTTA 1204047  
Qy 421 TTCAACGCAATGCGGTAGCCCTTTGAAAGTCTTAAGGTAATCAAGCATGSCATCCATGGG 480  
Db 1204046 TTGAAAACATATGATCTAACCTTTGAAATGAGCAAGTGGGTAGCAATTTGAATCTAAAT 1203987  
Qy 481 AAGATTGGCACTGTGCTTAAGGCAAGTGAATAAGCATTGAGTATGCTTGCC 540  
Db 1203986 AATACCTATAGCTGTGCAATGCTGTGCTGCAAAATGCTTCATTAATATGCGAC 1203927  
Qy 541 TTTGATTTTAAATGATGATTTGTTAAAGTTTGTTTTGTGATGATATAGCATAC 600  
Db 1203926 AACCAATGATGCTTCTTCAAGGTAAATATCACTTATCTTTTACATTAATATTTGCTT 1203867  
Qy 601 ATCACCATCATCAATTAATAATAAATCTGC-CGTTTGTGCTAAGTAACTATTTGTTG 659  
Db 1203866 GCATTTTCAAGTTTATAGCAATAATCCGCACTAATTTCCGAAAGAAACGGTATGTT 1203807  
Qy 660 TGAATATGATGATTTTGGGGCAAGCTGTCAGTGTGTCAAGATTTGCTTGCTAAT 719  
Db 1203806 CGCTTATTTTGTGTTTCCCAAAAGTCAAGC----- 1203775  
Qy 720 TGGCATGCTTGGTATGATGATGCGGCTTTGGCAATTTTCAAAATATACCAATTTTGTG 779  
Db 1203774 ---TATCTTATTTTCAACACACTTACTTATGATTAATGATTAATTTTGTCTT 1203718  
Qy 780 CCAATATATCCACATGCAATGCTGCGCATCAAGATGATCAAGGAGAGATTTAAATTTG 839  
Db 1203717 GGGCATTAATCTTCTAATCAATATAGCATATGATCTTCAGTCACTTCAAGCAAG 1203658  
Qy 840 TTGCCCCCTGAGCTGCCAATTTGAGATATGCTCAAGCTGAAACTTGCACAGTCAAGCA 899  
Db 1203657 TCGCTGCGGAGCTTTTAAGCTATATGCTCAAGCTGAAAACCTGAAGCTTATGTA 1203598  
Qy 900 CCACCAATTCATATTTTGGATGTTAATATTTCAAGTCAAGGCGTGCATATTTACCGC 959  
Db 1203597 CATTAAGTTCACATCTTCAATTC--AACATGACAAAGCGGAATCCCAATTTTCCGC 1203541  
Qy 960 CGACACCAACATCATGCTGCAATGTTTGGCATCTGCGCTACTAATGTGTAACAGTGC 1019  
Db 1203540 CCAATCAACTTTCAACACGAGGCTTTCGCCATTTCAATAACTAAGATTAACGCTATC 1203481  
Qy 1020 TTTTGGCATTTGAGCCTGTGATGCGATATTTGTGTGTGTTGTGCTTGAATGATCTC 1079  
Db 1203480 TTTTACCATTTTGAACCTGTAAATCCCAATTTGCTTGTGCTGCGGCA----- 1203429  
Qy 1080 GTGCTTGAATGATCAATGAACAATTTGAACATGCTGATCAACAGAAATGCTTGAAGCTT 1139  
Db 1203428 -----GAATTAATTCATATTCGCCATTTACTTCCACTCCCGCTTTAA 1203388  
Qy 1140 TGGGCGCAATATCTGCGCGTTCTTGGTTAATGCGGGGCTGATGATCTGATCGG 1199  
Db 1203387 GTGCGGTTTGAATTTCTGTGTGTTTAAACGCAAGCCCTGGGCTAATTAACATCATATCCG 1203328  
Qy 1200 CTGAGCTTAAGATTTACTGTCCAAAGCTGCCAAAATGAACAATGAACGCTGTGGGAGTT 1259

Db 1203327 TTTCAGTAACCATTCCTGATTTTAACTACAGATGAGAGGATATTTTGAGAGATT 1203268  
Qy 1260 TATTCGCCAAGTATGATTTGGGATTTACATCGATGATCTTATTTGTACCGTGATTGA 1319  
Db 1203267 TATCAATACAGATGATTTTTCGATATCAATACACGAAATTTAGCCCTGTGGGATA 1203208  
Qy 1320 CCAAAAAATTCACAGCTGACAGTCCAGAGTCCCAACGAGACGCGATATTTT 1377  
Db 1203207 AGAGATATTCACACAGAAAGACCTGTTTTCAGAGCCGATATGTAATATTTT 1203150

RESULT 5  
US-09-643-990A-1/c  
; Sequence 1, Application US/09643990A  
; Patent No. 6528289  
; GENERAL INFORMATION:  
; APPLICANT: Robert D. Fleischmann  
; Mark D. Adams  
; Owen White  
; Hamilton O. Smith  
; J. Craig Venter  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville,  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3 1/2 inch diskette  
; COMPUTER: Dell Pentium  
; OPERATING SYSTEM: MS DOS v6.22  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/643,990A  
; FILING DATE: 23-Aug-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,429  
; FILING DATE: 1995-06-07  
; APPLICATION NUMBER: 08/426,787  
; FILING DATE: 1995-04-21  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kenley K. Hoover  
; REGISTRATION NUMBER: 40,302  
; REFERENCE/DOCKET NUMBER: PB186P1C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-610-5790  
; TELEFAX: 310-309-8439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1830121 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1

Query Match 11.6%; Score 160.4; DB 3; Length 1830121;  
Best Local Similarity 49.6%; Pred. No. 9.6e-37;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy 1 TAAAGCTGAACCAATGTTAACTATACCTGCTCATTTATATTTTAAATGATC 60  
Db 1204439 TAAACATTTGAGTTAAATGCGTAATTTCTTCCGCGCTTTTCAAAAAGAACTATC 1204380  
Qy 61 CATGCTAGCAACAGCAGGTGACAAAGACAGCCCTTGAAGCGTGTGACTGCGCAA 120  
Db 1204379 GAGACTTGACAAAGCAGCGCAATTAATATCA-----TATCTCGCTTTGCAA 1204332

Qy 121 TACTGCCCATCAAGGTACCAACCTGATATATTTAGATGCCAATGTGCATACCTT 180  
Db 1204331 TGTGGCGGTAAATAATCTATGCTTGTTCATTTATGCAACA-----ATAACTTTG 1204278  
Qy 181 GAGTAAATCTTGTCTCAATTAATCGGGGATCTTACCGATTAATAACATATACCA 240  
Db 1204277 CGATGAAAAATTTTGCAGAGCGCCACCATCTGCACAAACAAATTAATATGTGG 1204218  
Qy 241 CTGCATACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCACCAAT 300  
Db 1204217 TTGATTAATTAATTCAGCTAATTCGAAAAATCAGCCCTTTTCCTGCTCGCTAGCAA 1204158  
Qy 301 CAATGCCAGATGATATGCGCATTAACCGGCCAAGCCCATCAATGTCTGATGTA 360  
Db 1204157 CAATGCAATTAACCTCAATATA-----AAGCCGACCAATGACCAACTGACT 1204107  
Qy 361 GCCGATTTGTACCTTTTGAATCATTAATAATATGCTCTATCAATATCATCAATATA 420  
Db 1204106 CCCCACTTTGTGCTTTAGAGTCATTAATCCAAAGCCATTAAGCTTGATGCACTAA 1204047  
Qy 421 TTCACAGCAGTGGGTAGCCCTTTGAAAGCTTAAGGATATCAAGCATGSCATCCATGG 480  
Db 1204046 TTGAAAAAGATGATCTAACCTTTGAAATGACGAAGTGCAGTATGAATCTAAT 1203987  
Qy 481 AAGATTGCGAGCTGTGCTTAAGGCAAGTACAGATTAAGCATTAAGTATGCTTGGC 540  
Db 1203986 AATPACTATGCTTGTGCGCAATCTGTGCTGCAAAATGTTATTAATTAATGCGCAC 1203927  
Qy 541 TTTGATTTTAAATGATGATGATGTAATAAGTTGTTTGTGATATATATGCAATACC 600  
Db 1203926 AACCAATGATGCTTCTTCAACAAGTAAATCACTTATCTTTACATTAATATGCTT 1203867  
Qy 601 ATACCATCATCAATTAATAAATCTGC-GCTTGTGCTTAAGTATGCTTGTG 659  
Db 1203866 GCCATTTTCAAGTTTGTAGCCAAATATCCGCACTATTTCCGCAAAAGAAACGATGCTT 1203807  
Qy 660 TGAATATGCTGATTTGGGGCAACGCTGTGCTGATGTGTCAGACATGCTTGGCTAAT 719  
Db 1203806 CGCTGATTTTCTTTTCCCAAAAGTCAGCC----- 1203775  
Qy 720 TGCCATCGTCTGTGATGATGATGCGGTTTGCAATTTTCAAAAATACGCAATTTTGTG 779  
Db 1203774 ---TATCTTATTTGTTCAACACACTTACTTACATTAATGATTAATGCGTAATTTGCTT 1203718  
Qy 780 CCAATTAATCCACCATGCGCATGCTGCCGATCAAGATGATCAGCGAGAGATTTAAATTG 839  
Db 1203717 GCGATTAATCTTCTAATCAATCAATPAGCGATCCATGATCTTCAGTCACGTTCAAGCAG 1203658  
Qy 840 TTGCCCCCTGAGTGCCAATTTGAGATATGCTCAAGCTGAAAACCTTGACAGCTCAAGCA 899  
Db 1203657 TCGCTGCGCAGCTTTTAAAGCTAATAGTGTCTCAAGCTGAAAACCTGAAAGCTTAAGTA 1203598  
Qy 900 CCACCAATTCATATTTTGGATGTTAATTAATCAAGTGCAGCGCTGCCAATTTACCGC 959  
Db 1203597 CATTAAGTTCAATCTTCAATTC---AACATGACAAAGGGGAATCCCAATTTTCCGC 1203541  
Qy 960 CGACACCAATCATGCTGTCATGTTTGGCATCTGCGCTACTAATGTCTGAACAGTGC 1019  
Db 1203540 CCATACCAACTTTCACACGAGCAGCTTTCGCAATTCATTAAGTAAAGTATGTAACGCTAC 1203481  
Qy 1020 TTTTGGATTTGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
Db 1203480 TTTTACCATTTGAACCTGTAATCCCAATTTGCTTGTGCTGCGCGCA----- 1203429  
Qy 1080 GTGCTTGAATGATCAATGAACATTTGAACATGCTGATCAGAGATGCTTGAAGCTT 1139  
Db 1203428 -----GAATTAATTCATATGCGGATTAATTCACACTCCGCTTTAA 1203388  
Qy 1140 TGGCGGCAATATCTGCGGCTTCTTGGGTTAATGCGGGGCTGATGATGATGATGATGATGAT 1199  
Db 1203387 GTGCGTTGTAATTTCTGCTGTTTAAACGCAAGCCCTGCGCTTAATTAACATCATATCTGC 1203328

```

Oy      1200   CTGAGCTTAAAGATTTCACTGTCGCAAGCTGCGCAAAATGCATGAAACGCGCTGTGGCAGTT 1259
Db      1203327 TTTCACAGTACCACTTCCTCGATTATTAACCTACAGTGTGAAGAGGATATTTTGAGGAAGTT 1203268
Oy      1260   TATCTGCCAAGGTAGATTTGGGATTACCATGATGATGCTAACTTTGTGACCGTGATTGA 1319
Db      1203267 TATCAATATCCAGTATGAGATTTTTTTTCAGATATCAATACACGGAATATTAAGCTGTGGGATTA 1203208
Oy      1320   CCAAAAAATTCACACGCTGACAGTCCAGAGCTGCGCAACCGACGACGCGCATATATTTT 1377
Db      1203207 AGAGATATATCCACACAGAAAGAAAGCCTGTTTGGCAAGCCCAATGATGATTAATATTTT 1203150

RESULT 6
US-10-158-865-1/c
: Sequence 1, Application US/10158865
: Patent No. 6846651
: GENERAL INFORMATION:
: APPLICANT: Fleischmann et al.
: TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
: Patent No. 6846651
: TITLE OF INVENTION: Thereof, and Uses Thereof
: FILE REFERENCE: PB186P2C1D1
: CURRENT APPLICATION NUMBER: US/10/158,865
: PRIOR APPLICATION NUMBER: US 09/557,884
: PRIOR FILING DATE: 2002-06-03
: PRIOR APPLICATION NUMBER: US 2000-04-25
: PRIOR FILING DATE: US 08/476,102
: PRIOR APPLICATION NUMBER: US 1995-06-07
: PRIOR FILING DATE: 1995-04-21
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1830121
: TYPE: DNA
: ORGANISM: Haemophilus influenzae
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (4747)..(4747)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9921)..(9921)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (10150)..(10150)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (29298)..(29298)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (36543)..(36543)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (36551)..(36551)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (36636)..(36636)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (40808)..(40810)
: OTHER INFORMATION: n equals a,t,c, or g
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (44416)..(44416)
: OTHER INFORMATION: n equals a,t,c, or g

```

|    |   |  |
|----|---|--|
| 1  | FEATURE:                                |  |
| 2  | NAME/KEY: misc_feature                  |  |
| 3  | LOCATION: (44905) .. (44905)            |  |
| 4  | OTHER INFORMATION: n equals a,t,c, or g |  |
| 5  | FEATURE:                                |  |
| 6  | NAME/KEY: misc_feature                  |  |
| 7  | LOCATION: (44975) .. (44975)            |  |
| 8  | OTHER INFORMATION: n equals a,t,c, or g |  |
| 9  | FEATURE:                                |  |
| 10 | NAME/KEY: misc_feature                  |  |
| 11 | LOCATION: (44593) .. (44593)            |  |
| 12 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 13 | FEATURE:                                |  |
| 14 | NAME/KEY: misc_feature                  |  |
| 15 | LOCATION: (45732) .. (45732)            |  |
| 16 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 17 | FEATURE:                                |  |
| 18 | NAME/KEY: misc_feature                  |  |
| 19 | LOCATION: (47036) .. (47036)            |  |
| 20 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 21 | FEATURE:                                |  |
| 22 | NAME/KEY: misc_feature                  |  |
| 23 | LOCATION: (51334) .. (51334)            |  |
| 24 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 25 | FEATURE:                                |  |
| 26 | NAME/KEY: misc_feature                  |  |
| 27 | LOCATION: (51602) .. (51602)            |  |
| 28 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 29 | FEATURE:                                |  |
| 30 | NAME/KEY: misc_feature                  |  |
| 31 | LOCATION: (51786) .. (51786)            |  |
| 32 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 33 | FEATURE:                                |  |
| 34 | NAME/KEY: misc_feature                  |  |
| 35 | LOCATION: (51805) .. (51805)            |  |
| 36 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 37 | FEATURE:                                |  |
| 38 | NAME/KEY: misc_feature                  |  |
| 39 | LOCATION: (55369) .. (55369)            |  |
| 40 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 41 | FEATURE:                                |  |
| 42 | NAME/KEY: misc_feature                  |  |
| 43 | LOCATION: (65309) .. (65309)            |  |
| 44 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 45 | FEATURE:                                |  |
| 46 | NAME/KEY: misc_feature                  |  |
| 47 | LOCATION: (65024) .. (65024)            |  |
| 48 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 49 | FEATURE:                                |  |
| 50 | NAME/KEY: misc_feature                  |  |
| 51 | LOCATION: (100691) .. (100691)          |  |
| 52 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 53 | FEATURE:                                |  |
| 54 | NAME/KEY: misc_feature                  |  |
| 55 | LOCATION: (102696) .. (102696)          |  |
| 56 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 57 | FEATURE:                                |  |
| 58 | NAME/KEY: misc_feature                  |  |
| 59 | LOCATION: (105121) .. (105121)          |  |
| 60 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 61 | FEATURE:                                |  |
| 62 | NAME/KEY: misc_feature                  |  |
| 63 | LOCATION: (107248) .. (107248)          |  |
| 64 | OTHER INFORMATION: n equals a,t,c, or g |  |
| 65 | FEATURE:                                |  |
| 66 | NAME/KEY: misc_feature                  |  |
| 67 | LOCATION: (117136) .. (117136)          |  |
| 68 | OTHER INFORMATION: n equals a,t,c, or g |  |

```

NAME/KEY: misc_feature
LOCATION: (119750) ..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924) ..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038) ..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344) ..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336) ..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) ..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) ..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) ..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) ..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) ..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) ..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) ..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) ..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) ..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) ..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

```

```

Query Match 11.6%; Score 160.4; DB 3; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 9.6e-37;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```

```

QY 1 TAAAGGTGACCAATGTTAAACAACCTTATCACCTGGCTGATTAATTTTAAATTGATC 60
DB 1204439 TAAACATTGAGCTAAATGCGTAAATTTCTGGCGGCTTTTCAAAAGAACCACTGATC 1204380
QY 61 CATGCTAGCACACAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTGACTGCTGCA 120
DB 1204379 GAACTTGGACAAAGAGGGGACAAATATACA-----TATCTCGCTTGGCA 1204332
QY 121 TACTGCCGATCAAGTGTACCAACTGATATATTTAGATGCCAAATGTGATCACCCTT 180
DB 1204331 TGTGGGCGTAAATAATTTATGCTTGTTCATGTATGACAA-----ATACCTTGG 1204278
QY 181 GAGTAAATCTGTGCTCAATTAATGCGGGCATCTTGACCGATTAATATACACTATCACCA 240
DB 1204277 CGATGAAAATTTTGCACAGCGCCACACTCTGACCAAAACAAATATATATGTGTG 1204218
QY 241 CTGCATPAACATATGCTGACAGCTCACTAAATCTTGAACCTTACCAAGCCACCAAAAT 300
DB 1204217 TTGATTAATTAATTCAGCTAATTCGAAAATCAGCCCTTTCCGTCCTCGCTAGCA 1204158
QY 301 CAATGCCAGTGATGATATGCAATPAAACCGTCCAAAGCCCATCAATTGCTGCATGTTGA 360
DB 1204157 CAATGCAATTTAACCTCAATATA-----AGCCCAAGCCCAATGCAGCAACTGTACT 1204107
QY 361 GCGGATGTTTGTACCTTTGAATCATTAATAATATGCTTGTATCAATATCATCAATATA 420
DB 1204106 CCCACATTTTGTGCTTTTGAAGTCAATTAATCCACAGAAATGCTTGTGATGACTTA 1204047
QY 421 TTCACAGCGATGCGGTAGCCCTTTGAAGTCTTAAAGTATCAAGATGAGCATCCATGGG 480
DB 1204046 TTGAAAACGATGATCTAACCCCTTTGAAATGACGAATGGCGGTGCAATGAATCTAAAT 1203987
QY 481 AAGATTGCAAGCTGTGCTTAAAGCAAGTCAATTAAGCATTAAGTATGCTTGGC 540
DB 1203986 AATACCTATAGCTGTGCAATGCTGTCTGCCAAATGTTCAATTAATTAAGGCGACC 1203927
QY 541 TTTGATTTTAAATGATGATTTGGTAAAGTTTGTTTTGAATGATATATATGCAATACC 600
DB 1203926 AACCAATGATGCTTCTTCAACAGGTAAATCACTTATCTTTTACCAATTAATTTGCTT 1203867
QY 601 ATCACCATCATCAATAATTAATAAATCTGC-CGTTTGGTGGCTAATGATTAATTTGTTG 659
DB 1203866 GCCATTTTCAGTTTATAGCAATTAATCCGACATATTTCCGAAAAGAAACGGTATGTTT 1203807
QY 660 TGAATTAATGATGATTTGGGCAACGCTGTCACTGTGCTCAAGCATTTGCTTGGCTAAT 719
DB 1203806 CGCTGATTTTGTGTTTCCCAAAAGTCAAGC----- 1203775
QY 720 TGGCATGCTGTGATGATGATGGCGTTTGGCAATTTTCAAAATATGCAATTTTGTG 779
DB 1203774 ---TATCTCATTTGTTCACACACTTACTTTAGCATTTATGATTAATGCGTAATTTGCTT 1203718
QY 780 CCNAATATCAACCATGCAATCGTCCGATCAAGATGATCAGCGAGATTTTAAATTG 839
DB 1203717 GGGGATATCTTTAATCATATATAGCATTCATATGATCTTCAATGCAAGTCAAGCAG 1203658
QY 840 TTGCCCCCTTGAAGTGCCTTGAATATGCTCAAGCTGAAAACCTTGACAGCTCAAGCA 899
DB 1203657 TCGCTGCCGAGCTTTTAAAGCTATATAGTGTCTCAAGCTGAAAACCTTGAGTA 1203598
QY 900 CCACCAAAATCAATTTTGGATGTTAATTAATTCAGTGCAGGCGGCAATTTTACCGC 959
DB 1203597 CATTAAGTTCAACATTTTCAATTC---ACAAATGACAAAGGGAATCCCAATTTTCCGC 1203541
QY 960 CGACACCAACAAATCATGCTGCAATGTTTGGCATCTGCGCTTAAATGTCGTAACAGTGC 1019
DB 1203540 CATTAACCACTTTCAACACAGGAGCTTGGCATTTTCAATTAATTAAGTATGCGTATC 1203481
QY 1020 TTTTGGCATTTGAGCTGTGATGAGCATATATGATGATTTTGTGCTTGAAGTATGATCTC 1079
DB 1203480 TTTTACCATTTGAACTGTATATCCCACAATAGGCTTTGTGCTGCGCGCA----- 1203429

```





Db 941 AGGCGGAAAAAGCTTACAGCGCCGAGATGCGTGAACGCGAGCGGACCGCATGG 882  
Qy 497 CCTAAGCAAGTGCAGATAAGGCAATGAGTATGCTTGGCTTGAATTTAATGA 556  
Db 881 CCCAGCCCGAGCGGAGCGGCGGTGAAATAGTTGGGCGCCAGCAATCTTCAAGTTC 822  
Qy 557 TGAATGTAATAAGTTGTTTTTGTATGATATATGCAATCCATCAGCATCAAT 616  
Db 821 CCAACCGGAGAGCTGTGAACTGGAAGCCAGCACTTCCGCGTCTTCTCATC 762  
Qy 617 AAATTAATAATCTGCCCTTTGGTGGCTAAGTAAGTATTTGTTGATTAATGTTGATTT 676  
Db 761 AGGCGGAAAGC-----CTTGAAGTCCGGCTTGTTCAGGCCGAA 724  
Qy 677 GGGGCAACGCTTGTCACTGTGTCAAGCATTTGCTGCTAATTTGGCATGCTTGGTAG 736  
Db 723 CGACCGACAGCGACGGTATCGGATCAGCGTGGGTCA--GGGCAATCGGCGCGATTC 666  
Qy 737 ATGATGCGGTTTGGCAATTTTCAAAAATAGCAATTTTGTGCAATTAATCCAGCATG 796  
Db 665 ACCACGACCTGGGGGAGACCGCGGAGATCCGGTCTTGGCGAGGTGTAGTCAAGCATG 606  
Qy 797 CCAATGCTGCGATCAAGATATACGCGAGAGATTTAAATTTGTCCTTGAAGTCC 856  
Db 605 CCGTCTGAGCGATCAATATGCTTTCCTGACGTTAGCAAGCGGTCCACCTCGGCGTTG 546  
Qy 857 AAATTTGAGATATGCTCAAGCTGAATACTTGAACGCTCAAGCAACCAATCCATTTT 916  
Db 545 AGCGATCGAGGTTTTCAGCTGAAGCTGACAGCTTCAACAGTACACTCGATGTCG 486  
Qy 917 TGGATCGTTAATTAATCAAGTGAAGCGGTCCAAATATTAACCGGACACCAACATCATG 976  
Db 485 TCG---GCCAGCAGGTGAGCGCGGGTCCGAGGTTGCCCGGAGCGGACGCTTG 429  
Qy 977 CTTGCAATGTTTGGCATCTCGCCCTACTAATGTCTGAACAGTGTCTTTGGCATTTGAGCTT 1036  
Db 428 TCCGCGGCGCACCCCAATTTGCCACACGAGGTGGTCAAGGTGCTTCCGCTTGAACCG 369  
Qy 1037 GTGATGCGCATGATGTTGGTGTGTTG 1061  
Db 368 GTGATGCGCAGATCGGGGCTTCCG 344

RESULT 9  
US-09-252-991A-7861  
; Sequence 7861, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7861  
; LENGTH: 1401  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7861

Query Match 8.5%; Score 117; DB 3; Length 1401;  
Best Local Similarity 48.4%; Pred. No. 3.8e-25;  
Matches 506; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

Qy 17 TTAAACAACCTTACCTCGCTCATATATTTTAAATGATCCATGCTAGCACAAGCA 76  
Db 14 TTGGCGCAAGCGGCTCGCTCTTCTTCAAGATCTTGAACATGTCAGAGCTCGCGAGGCC 73

Qy 77 GTGACAAAAGCAGACGCTTAGCGGTCTTTGACTGGCTCCAAATACCTGCCATCAAGT 136  
Db 74 GGGGACAAAGCAGACCGCATGCGCTTCCG-----GGGCGAGCTGGGGCTGGCG 124  
Qy 137 GTACCAACTGATATATTTTAAATGATGCAATGTCATCACTTTGATTAATCTTGCTCA 196  
Db 125 ACTGCTTCGTCCAGCGTTGGAGCGGACACGAGGGTATCCCGTGTGCCAATGCTGGGCA 184  
Qy 197 AATATCGGGGCACTTGAACGATTAATAATACACTATCCACCACTGCAATCAATATGCT 256  
Db 185 ATGACCCGGCGTACGGCCAGACAGTACACCGCCCGGAGAAAGCGCGGAGCGGCTCG 244  
Qy 257 GACAGCTCACTAAATCTTACCTTTTACCAAGCCCAACCAAAATCAATGCAATGATAT 316  
Db 245 CGGAGTCAATGAAATCGGGCCCTTGGCCCTCTCCGCGGCGAGCAGCATCACTTCCG 304  
Qy 317 ATGCAATAAACGTCGCCAAGCCCATCAATTTGCTGCAATGTTGAGCCGATGTTGACT 376  
Db 305 TCGATGT---CGGACCCAGCCCTCGATCGCGGCGGAGGCGGCGCGAGCTTGGTGGC 361  
Qy 377 TTTGAATCAATTAATAATGCTTGTCTATCAATATCAATATATTTCAAGCGATGGCT 436  
Db 362 TTGGAATCGTGTATGATCACTACGCGCTCGCTCGGATCAACATGCGAGATGAGCC 421  
Qy 437 AGCCCTTTGAAAGTCTTAAGGATATCAAGCATGCAATCCATGGAAGATTGGCAGCTGTG 496  
Db 422 AGGCGGAAACGCTTTCAGCGCGCGGAGCATGCTCAACGCGAGCCGACCGCATGG 481  
Qy 497 CCTAAGCAAGTCAAGTAAGGATTAAGTATGATGTTATGCTTGTGATTTTAAATGA 556  
Db 482 CCCAGCGCCAGCGCGGAGCGCGGTGGAATATGTTGGCGGCGCACGAGATCTTCAAGTTC 541  
Qy 557 TGAATGTAATAAGTTTGTGTTTGTGATGATATTAAGCATATACATCAATCATCAAT 616  
Db 542 CCAACCGGAGCAGCTTGTGCAACTGAACTGAAACCGACCACTTGGCGCTTCTTCTGATC 601  
Qy 617 AAATTAATAATCTGCCCTTTGGTGGCTAAGTAAGTATTTGTTGATTAATGATGATTT 676  
Db 602 AGGCGGAAAGC-----CTTGAAGTCCGGCTTGTTCAGGCCGAA 639  
Qy 677 GGGGCAACGCTTGTCAAGTGTGTCAGCATTTGCTTGGCTAAATTTGCAATGCTTGTGATG 736  
Db 640 CGACCAAGCAGCGCAGCGTATCGGCGATCAGCGGTCCGGGTCA--GGGCAATGCGCGCATTC 697  
Qy 737 ATGATGCGGTTTGGCAATTTTCAAAAATACCAATTTTGTGTCCTCAATATCAAGCATG 796  
Db 698 ACCACGACCTGGGGGAGACCGCGAAGATCCGGTGTGCGCAGGTGTAGTCAAGCCATG 757  
Qy 797 CCAATGCGCGATCAAGATGATCAGCGGAGAGATTTAAATTTGTCGCCCTTGAAGCTGCC 856  
Db 758 CCGTCTGAGCGATTCATATGCTTCTTGTGAGCTTACGACGATCCGCACTCGGCGTTG 817  
Qy 857 AAATTTGAGATATGCTCAAGCTGAATACTTGAAGCTCAAGCAACCAATTCATATTT 916  
Db 818 AGCGATCGGAGTTTCCAGCTGGAAGCTTCGACAGCTTCCAAACAGTACAGCTGATGTCG 877  
Qy 917 TGAATGTTAATTAATTAAGTGAAGCGGTCCCAATTTTACCGCGGACCAACCAATCATG 976  
Db 878 TCG---GCCAGCAGGTGAGCGCGCGGGTGCACAGGTTGCCCGGAGCGGAGACGCTTG 934  
Qy 977 CTTGCAATGTTTGGCATCTCGCCCTACTAATATGTCTGAACAGTGTCTTTGGCATTTGAGCTT 1036  
Db 935 TCCGCGGCGCACCGCATTTTCCGCCACACAGGTTGTCAAGGTGCTTCCGCTTGAACCG 994  
Qy 1037 GTGATGCGCATGATGTTGGTGTGTTG 1061  
Db 995 GTGATGCGCAGATCGGGGCTTCCG 1019

RESULT 10  
US-09-701-229-1/c  
; Sequence 1, Application US/09701229  
; Patent No. 6890910





|                |      |   |      |
|----------------|------|---|------|
| D <sub>b</sub> | 399  | -----GCAGAACGTTTCGATTCGCGCAAGATCTCTACGCCACGGCTCG              | 357  |
| Q <sub>y</sub> | 1139 | TTGGCGCGCACAATCTCGCGCTTCTTGAGTTAATGCCGGGGCTGATGATGATCTGATCG   | 1198 |
| D <sub>b</sub> | 356  | GCCGCTCGCTGACGGAAGATGAGCGCAGCCCAATTTCCCGGACTGGGCGACATCAGGCTC  | 297  |
| Q <sub>y</sub> | 1199 | GCTGAGGTTAAGATTCCTACTGTCTCCAGCTGCCAAATATGACATACAGCCTGTGGGAGT  | 1258 |
| D <sub>b</sub> | 296  | GCCGCCAGCAGCCAGGTATCATTTGAGTCCGCGCAGTCAGTTCGACGATTTCCGGTAAAT  | 237  |
| Q <sub>y</sub> | 1259 | TTATTCGCCAAGTGAATGGGAAATTACATCGATGATGATCTTTGTACCGGTGANTG      | 1318 |
| D <sub>b</sub> | 236  | TTATTCACGCCCCGGAGGCGCGAAGCGGGGTTCATCAGCGCGGGGTTACGCCCGCGCC    | 177  |
| Q <sub>y</sub> | 1319 | ACCAAAAATTCACAGCTGACAGTTCACAGGCTGCCAAACGACGACGGCATTAATTTTTT   | 1377 |
|                | 176  | ATGAAAAAGTCCACGATGACAGGCGCGGTCTATGCCACGCCCAATTAATGACACATTTTTT | 118  |

RESULT 12  
US-09-790-988-1  
; Sequence 1, Application US/09790988

```

APPLICANT: SHIGEMOBU, SHUJI
APPLICANT: MATANABE, HIDEMI
APPLICANT: HATTORI, MASAHISA
APPLICANT: SADAKI, YOSHIFYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

```

|                      |              |                   |                |               |
|----------------------|--------------|-------------------|----------------|---------------|
| Query Match          | 6.1%         | Score 84.8        | DB 3           | Length 640681 |
| Best Local Similarly | 50.2%        | Pred. No. 8.4e-14 |                |               |
| Matches 289          | Conservative | 0                 | Mismatches 257 | Indels 30     |
|                      |              |                   |                | Gaps 2        |

|    |        |   |        |
|----|--------|---|--------|
| QY | 802    | GTGGCGAATCAAGTATCATACGGGAGAGATTAAATTTTGGCCCTTGAGGCTGCCAAATT       | 861    |
| Db | 240727 | GTATTCAGTATGATATATCTTCCTTAATTTAAGAAATTAAGTATTTAAATT               | 240786 |
| QY | 862    | TGAGATATGCTCMAAGCTGAAAACTTGACAGCTCAAGCACCAACAAATTCATATTTTGAT      | 921    |
| Db | 240787 | AAAGTATTTTCTACTGTGAAAACTAGACAGTTCTATATATGATATAATTCGCTTCTT---      | 240843 |
| QY | 922    | CGTTAATATTTCAAGTCGACGGCGCTGCCAATATTACCGCGGACACCAACAATATGCTGCG     | 981    |
| Db | 240844 | ATCAAGTATTTCTACTAGCGGGAACACTATATTTACCACTTGAAAAAGCTTTATATCTGA      | 240903 |
| QY | 982    | ATGATTTTSCCATCTGGCCTACTAATATGCTTAAACAGTCTTTTGGCACTTGAGGCTTGAT     | 1041   |
| Db | 240904 | TTTTTTTGCATTTTTTTTCAATATATGTAAGTACTACAGTACTTTTTTCCATATGATCCCTTTAT | 240963 |
| QY | 1042   | GGCGATGATGCTGTGTTTGTTTGCTGTAAGTAAATCTCGTCTTTGAGTATCAATGAA         | 1101   |
| Db | 240964 | TGAATATATATAGACACAGTCTACTTCTCTAG-----AAAA                         | 240996 |
| QY | 1102   | CAATTGAACATGCGGATCTACAGGAATGCTTGAGCTTTGGCGGACCAATCTGGCGCT         | 1161   |
| Db | 240997 | TAGTTCATATCTACTAATATATCTCAATACTACGAAACGTGCTTTAATTTAAATAGGTTT      | 241056 |
| QY | 1162   | TCTTGGGTTAATGCCGGGCTGATGATGATCTGATGCGCTGAGCTTAAAGATTCACTGTTC      | 1221   |

|    |        |  |        |
|----|--------|--|--------|
| Db | 241057 | AAATGAGAAATATCCCGGACTTTATACAAATTAACTGTGATTTAGAAATCCACTGATGTC | 241116 |
| Qy | 1222   | CAAGTCGCCAAATGACACATGAGACCGCTGTGGCGATTATCTGCCAAGTAGATGGG     | 1281   |
| Db | 241117 | TAACTTCTTAAACATATTTGCAATTTTGGAGATTTTAAATAAATTTGAAGGATGTT     | 241176 |
| Qy | 1282   | ATTACCATGATGATGCTTAACTTTGTGACCGTGATTGACCAAAAAATTACACGTCGACG  | 1341   |
| Db | 241177 | AGATTCATCATTAATTTTAAAGTTTATTCCTTTTAAAAAAAATTATATCAAGATAT     | 241236 |
| Qy | 1342   | TCCAGAGCTGCCAAACCGACGACGCAATTTTT                             | 1377   |
| Db | 241237 | TTCGTGTAACCCATTTCTTAATAATTAATTTTTT                           | 241272 |

RESULT 13  
US-09-107-532A-1574/c  
; Sequence 1574, Application US/09107532A  
; Patent No. 6583275

; GENERAL INFORMATION:

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 7310  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02354

```

?      CURRENT READABLE FORM:
?      MEDIUM TYPE: CD-ROM ISO9660
?      COMPUTER: PC
?      OPERATING SYSTEM: <Unknown>
?      SOFTWARE: ASCII
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/09/107, 532A
?      FILING DATE: 30-Jun-1998

```

APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Arimbellio, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 893-5007  
TELEFAX: (781) 893-8277  
INFORMATION FOR SEQ ID NO: 1574:  
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 1368 base pairs
;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)

```

```
;
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;       ORGANISM: Enterococcus faecium
;   FEATURE:
```

```

; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1368
; SEQUENCE DESCRIPTION: SEQ ID NO: 1574::
US-09-107-532A-1574

```

|                       |  |                  |                 |              |
|-----------------------|--|------------------|-----------------|--------------|
| Query Match           | 5.0%   | Score 83;        | DB 3;           | Length 1368; |
| Best Local Similarity | 53.6%  | Pred. No. 1e-14; |                 |              |
| Matches 173;          | Conservative   | 0;               | Mismatches 150; | Indels 0;    |
|                       |  |                  | Gaps            | 0;           |
| 312                   | ATGATATATCCCAACCGTCGCAAGCCCATCAATTTGCTGCATAGTGGAGCCCGATGTTTG | 371              |                 |              |

312 ATGATATGCCATAAACCGTGCCAGCCCATCAATTGCTGCAATGTTGAGCCGATGTTG 371

```

Db      1069 ATACAGACCTTAGACGTTTCAATCCACTTAGAGCCATTGATGCTAAATATG 1010
Qy      372 TACCTTTGAATCAATTAATATGCTGCTATCATATCATATATATTCACAGCAT 431
Db      1009 TTGCTTTGAATATATATAAATTTTCTGCTTGATTTCTCTACATATATGGGAT 950
Qy      432 GCGGTAGCCCTTGAAGTCTTAAGGTATCAAGCATGAGCATGCATGGAAGATGGCAG 491
Db      949 GTGGTAGCCATGAAAAAATGCAAGCTTTCTGTATAGCTTCAATTGATATACATACA 890
Qy      492 CTGTGCTTAAGGCAAGTGCAGATTAAGCATTAAGTATAGCTTGCTTTGATTTTGA 551
Db      889 ATTTGCTACAGAAATAGCAGCTAGTGCATTTTCTATCATATATGCTTACGTAGCCAA 830
Qy      552 ATTGATGATGATGTAATAAGTTGTTTGTGATGATATATATGCATACATACATCAT 611
Db      829 GTTCAGATATTTCCATTAATTTTCTTTTATATATATGATATGATATGTCAGCATACA 770
Qy      612 CAATATAATTAATATGCGCTT 634
Db      769 CACCGCTTCAACACTTCTTTT 747

```

## RESULT 14

```

US-09-252-991A-7787
; Sequence 7787, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196, 136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7787
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7787

```

Query Match 5.1%; Score 69.8; DB 3; Length 567;

Best Local Similarity 55.0%; Pred. No. 6.9e-11;

Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

Qy      327 CCGTGCCAGCCCATCAATGCTGCAATGTTGAGCCGATGTTGTACTTTGAAATCAT 386
Db      262 CGGCACCCACCCCTTCATGCGCCAGGCGGCGGCGGCGCTTGGAATCGT 321
Qy      387 TAAATATGCTTGTCTATCAATATCATCAATATATATATATATATATATATATAT 446
Db      322 CGTAGTAGCTCAGCCCTGCGCGCTGCGGTACCACTGCGAGCATATGACAGCGGAAA 381
Qy      447 AATCTTAAGGATATCAAGCATGCGATCCATGGAAGATTTGGACGCTGTGCTTAAGCAA 506
Db      382 ACCCTTTCACGCGCCAGCATGCGCGTGAACGCGACGCGCATGCGCCGACGCCCA 441
Qy      507 GTGCAGATTAAGGATTAAGTATGATGCTTGTGCTTGAATTTTAAATGATGATGTA 566
Db      442 GCGCGGCGAGCGCTTGAATATGTTGGGCGCCAGCATCTTCAATTCGCCAACGCGCA 501
Qy      567 AAAGTTTGT 575
Db      502 GCAGCTTGT 510

```

## RESULT 15

US-09-221-017B-881/c

```

; Sequence 881, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221, 017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Montoy, Gladys H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 881:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...2402
US-09-221-017B-881

```

Query Match 4.9%; Score 67; DB 3; Length 2402;

Best Local Similarity 46.6%; Pred. No. 1.1e-09;

Matches 334; Conservative 0; Mismatches 365; Indels 18; Gaps 3;

```

Qy      355 GGTGAGCCGATGTTGTAACCTTTGATCATTAATAATATGCTTGTATCATATCATC 414
Db      2165 GGTAGATTTTACGTTGTAGCTTTGAGTCAATTAATCAACACCTTTACGCGAGC 2106
Qy      415 AATATATTCACAGCATGCGGTAGCCCTTTGAAGTCTTAAGGATATCAAGATGCGATC 474
Db      2105 AATCTTTTCAGTCGATGCGGTACATCTTGAATCTCGCAAGGCTTCTCGAATAGCTC 2046
Qy      475 CATGGGAAGATTGGCAGCTGTGCTTAAGGCAAGTGCAGATTAAGCATTAAGTATG 534
Db      2045 ATTCTTAATATCATAGCTTTGGCAGCAATGAGCATGTCATAGCATTTGTGCTGTG 1986

```

|    |      |  |      |
|----|------|--|------|
| Oy | 535  | CTTGCCCTTGAATTTTAAATGATGGAATGGTAAAGTTGTTTTGATGATATGAAATGC        | 594  |
| Db | 1985 | CATGCCGACAGAGCTAAAGTTCTCATCTCAATACAAAGGTGAATTAAGATTCTATGAC       | 1926 |
| Oy | 595  | CATACCATCACCATCATCAATTAATAAAAACTCCGGTTGGTGGCTAAAGTAACTATT        | 654  |
| Db | 1925 | AACCAATTCTGCTTTTTCATTGATTCGAAGCAGTTGTAATTATCCCGTTCGAGCTTCATATGC  | 1866 |
| Oy | 655  | TGTTGTGATATATGATGATGATTTGGGGGCAAGCCTTGTAAGTGTGTGCAAGCATTTGCTTGGC | 714  |
| Db | 1865 | GAAAGCGTGAAGAC-----GGGCTACAGAGATGTTCCGCTACCCATCCGCTGAT           | 1815 |
| Oy | 715  | TAAATTGGCATCGTCTTGGTAGATGATG-----GGGGTTGGCAATTTTCAAAAATAGC       | 768  |
| Db | 1814 | AAAAGATCATCTCCCGTAGATGAAGCAATCTCCGGGTGCTGATTTCCAGATGATTC         | 1755 |
| Oy | 769  | CAATTT---TTGTGCAAAATATATCACCATCCATCCGCGCCGATCAAGATGATCAAGGGA     | 825  |
| Db | 1754 | CATTTTGCCTTCTGTGATTAAGCTGGAATCGGTGATTCGATGATCTTAATATATCCGTGT     | 1695 |
| Oy | 826  | GAGATTTTAAATTGTGTGCCCTTGAGCTGCCAAATTGAGATATGCTCAAGCTGAAACT       | 885  |
| Db | 1694 | AATATTGAGCAGATAGGCCCATTAAGCCCTGAAGTCGTAATGTGTGTCACGTTGAAACT      | 1635 |
| Oy | 886  | TGACAGCTCAAGCACCACCAATCCATATTTTGGATCGTTAATAATTCAGTGAAGCGCT       | 945  |
| Db | 1634 | GCTCAACTCTATTAATATATTAAGATGAGATCGTAAGCGACTGTCTGGCCAGACTTAA       | 1575 |
| Oy | 946  | GCCAAATATTAACCGCGACACCAACATATCAGCTGCAATGTTTTCGCAATCTCGCCTACTAA   | 1005 |
| Db | 1574 | ACCGACATATGCTGTGCAAACTTACATCCAAACCGGCTTACAGAGCGTGTGATCAACCA      | 1515 |
| Oy | 1006 | TGTTGTATACATGCTTTTGGCATTTTGAGCCTGTGATAGCCATATATGTTGTGTTGT        | 1062 |
| Db | 1514 | CATGGTGTGTTTGTGCTTGGCATTACTACCCGTATAGCACCAATTAATGATCGTT          | 1458 |

Search completed: May 18, 2006, 09:34:35  
Job time : 268 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2006, 09:31:41 : Search time 1686 Seconds  
(without alignments)  
10057.495 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736

Perfect score: 1380

Sequence: 1 taaagcgcgaaccatgttaa.....aagcgcgcatacttttttga 1380

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues 27508734

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 22

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description                            |
|------------|-------|-------------|---------|----|--|
| 1          | 1380  | 100.0       | 1398    | 8  | US-10-282-122A-27195 Sequence 27195, A |
| 2          | 1380  | 100.0       | 96109   | 8  | US-10-672-787-35 Sequence 35, Appl1    |
| 3          | 231   | 16.7        | 1344    | 8  | US-10-282-122A-8940 Sequence 8940, Ap  |
| 4          | 160.4 | 11.6        | 1314    | 3  | US-09-815-242-7090 Sequence 7090, Ap   |
| 5          | 160.4 | 11.6        | 1314    | 8  | US-10-282-122A-22236 Sequence 22236, A |
| 6          | 160.4 | 11.6        | 1314    | 10 | US-10-958-215-400 Sequence 400, App    |
| 7          | 160.4 | 11.6        | 2925    | 15 | US-11-194-246-220 Sequence 220, App    |
| 8          | 160.4 | 11.6        | 11498   | 4  | US-09-754-468-44 Sequence 44, Appl1    |
| 9          | 160.4 | 11.6        | 1830121 | 9  | US-10-329-670-1 Sequence 1, Appl1      |
| 10         | 160.4 | 11.6        | 1830121 | 10 | US-10-981-687-1 Sequence 1, Appl1      |
| 11         | 160.4 | 11.6        | 1830121 | 10 | US-10-795-159-569 Sequence 569, App    |
| 12         | 154.4 | 11.2        | 14324   | 10 | US-10-795-159-569 Sequence 569, App    |
| 13         | 154.4 | 11.2        | 908766  | 10 | US-10-795-159-569 Sequence 569, App    |
| 14         | 152.8 | 11.1        | 1281    | 8  | US-10-282-122A-25165 Sequence 25165, A |
| 15         | 152.4 | 11.0        | 1314    | 10 | US-10-958-215-402 Sequence 402, App    |
| 16         | 142.2 | 10.3        | 1305    | 8  | US-10-282-122A-30713 Sequence 30713, A |
| 17         | 139.6 | 10.1        | 1317    | 3  | US-09-741-669-259 Sequence 259, App    |

|    |       |      |        |    |   |
|----|-------|------|--------|----|---|
| 18 | 139.6 | 10.1 | 1317   | 3  | US-09-815-242-5932 Sequence 5932, Ap    |
| 19 | 139.6 | 10.1 | 1317   | 4  | US-10-282-122A-20229 Sequence 20229, A  |
| 20 | 139.6 | 10.1 | 28277  | 4  | US-09-754-468-205 Sequence 2, Appl1     |
| 21 | 132.8 | 9.6  | 2850   | 15 | US-11-194-246-205 Sequence 205, App     |
| 22 | 132.8 | 9.6  | 5640   | 15 | US-11-194-246-205 Sequence 225, App     |
| 23 | 123.6 | 9.0  | 1308   | 8  | US-10-282-122A-32629 Sequence 32629, A  |
| 24 | 117.8 | 8.5  | 1182   | 8  | US-10-282-122A-33528 Sequence 33528, A  |
| 25 | 117.8 | 8.5  | 1347   | 3  | US-09-815-242-7912 Sequence 7912, Ap    |
| 26 | 117   | 8.5  | 1347   | 8  | US-10-282-122A-30468 Sequence 30468, A  |
| 27 | 113   | 8.2  | 1404   | 8  | US-10-282-122A-41274 Sequence 41274, A  |
| 28 | 111   | 8.0  | 1317   | 8  | US-10-282-122A-33003 Sequence 33003, A  |
| 29 | 110.8 | 8.0  | 4942   | 10 | US-10-450-763-25620 Sequence 25620, A   |
| 30 | 110.8 | 8.0  | 4944   | 10 | US-10-450-763-25620 Sequence 25077, A   |
| 31 | 110.8 | 8.0  | 4944   | 10 | US-10-450-763-25985 Sequence 25985, A   |
| 32 | 110   | 8.0  | 1350   | 8  | US-10-282-122A-31716 Sequence 31716, A  |
| 33 | 109.4 | 7.9  | 1317   | 3  | US-09-815-242-9978 Sequence 9978, Ap    |
| 34 | 109.4 | 7.9  | 1317   | 8  | US-10-282-122A-33835 Sequence 33835, A  |
| 35 | 103.8 | 7.5  | 1314   | 8  | US-10-282-122A-23400 Sequence 23400, A  |
| 36 | 94.2  | 6.8  | 1314   | 8  | US-10-282-122A-19656 Sequence 19656, A  |
| 37 | 92.8  | 6.7  | 2058   | 10 | US-10-450-763-8816 Sequence 8816, Ap    |
| 38 | 92.8  | 6.7  | 4831   | 10 | US-10-450-763-30346 Sequence 30346, A   |
| 39 | 92.6  | 6.7  | 1317   | 8  | US-10-282-122A-41740 Sequence 41740, Ap |
| 40 | 89.6  | 6.5  | 3117   | 10 | US-10-450-763-4899 Sequence 4899, Ap    |
| 41 | 89.6  | 6.5  | 3117   | 10 | US-10-450-763-25044 Sequence 25044, A   |
| 42 | 88.6  | 6.4  | 1353   | 8  | US-10-282-122A-17455 Sequence 17455, A  |
| 43 | 84.8  | 6.1  | 640681 | 3  | US-09-790-988-1 Sequence 1, Appl1       |
| 44 | 83    | 6.0  | 1359   | 8  | US-10-282-122A-21621 Sequence 21621, A  |
| 45 | 81    | 5.9  | 1071   | 8  | US-10-282-122A-37050 Sequence 37050, A  |

## ALIGNMENTS

RESULT 1  
US-10-282-122A-27195/c  
Sequence 27195, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308

```

; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PMLM
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 27195
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
; OS-10-282-122A-27195

```

|                            |        |             |      |             |
|----------------------------|--------|-------------|------|-------------|
| Query Match                | 100.0% | Score 1380  | DB 8 | Length 1398 |
| Best Local Similarity      | 100.0% | Pred. No. 0 |      |             |
| Matches 1380, Conservative | 0      | Mismatches  | 0    | Gaps 0      |

|    |      |  |     |
|----|------|--|-----|
| Qy | 1    | TAAGCGTGAACCAATGTTAAACAATCTATACCTGCTCATTAATTTTAAATTGATC      | 60  |
| Db | 1383 | TAAAGCGTGAACCAATGTTAAACAATCTATACCTGCTCATTAATTTTAAATTGATC     | 132 |
| Qy | 61   | CATGCTAGCAACAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCA   | 120 |
| Db | 1323 | CATGCTAGCAACAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCA   | 126 |
| Qy | 121  | TACGCGCCGATCAAGTGTACCAACCTGATATATTTTAAATGCAATGTCATCCCTT      | 180 |
| Db | 1263 | TACTGCGCCGATCAAGTGTACCAACCTGATATATTTTAAATGCAATGTCATCCCTT     | 120 |
| Qy | 181  | GAGTAAATCTTGCTCAATTAATCGGGGCACTTTGACCGATTTAAATACACTATCCACCA  | 240 |
| Db | 1203 | GAGTAAATCTTGCTCAATTAATCGGGGCACTTTGACCGATTTAAATACACTATCCACCA  | 114 |
| Qy | 241  | CTGATATAACATATGCTGACAGTCACTAAATCTTGACCTTTTACCAAGCCCAACCAAT   | 300 |
| Db | 1143 | CTGATATAACATATGCTGACAGTCACTAAATCTTGACCTTTTACCAAGCCCAACCAAT   | 108 |
| Qy | 301  | CAATGCCAGTATGATATGCCATAAACCGTGGCAAGCCCATCAATTGCTGCATAGTTGA   | 360 |
| Db | 1083 | CAATGCCAGTATGATATGCCATAAACCGTGGCAAGCCCATCAATTGCTGCATAGTTGA   | 102 |
| Qy | 361  | GCCGATGTTTGTACTTTTGAATCATTAATATATGCTTGTCTATCATATCATATATA     | 420 |
| Db | 1023 | GCCGATGTTTGTACTTTTGAATCATTAATATATGCTTGTCTATCATATCATATATA     | 96  |
| Qy | 421  | TTCAAGGAGATGCGGTAGCCCTTTAAAGTCTTAAGGGATTAACATAGGCATCCATAGG   | 480 |
| Db | 963  | TTCAAGGAGATGCGGTAGCCCTTTAAAGTCTTAAGGGATTAACATAGGCATCCATAGG   | 90  |
| Qy | 481  | AAGATTTGGCAGCTGTGCTTAAAGCAGTGCAGATTAAGCATTTAGTAGTTATGCTTGC   | 540 |
| Db | 903  | AAGATTTGGCAGCTGTGCTTAAAGCAGTGCAGATTAAGCATTTAGTAGTTATGCTTGC   | 84  |
| Qy | 541  | TTTGATTTTAAATGATGATTTGGTAAAGTGTGTTTTTGAATGATATTAATGCCATATC   | 600 |
| Db | 843  | TTTGATTTTAAATGATGATTTGGTAAAGTGTGTTTTTGAATGATATTAATGCCATATC   | 78  |
| Qy | 601  | ATCACCATCATCAATAATAATAAAAAATCGCGTTGGTGGCTAAAGCTATTTGTTGT     | 660 |
| Db | 783  | ATCACCATCATCAATAATAATAAAAAATCGCGTTGGTGGCTAAAGCTATTTGTTGT     | 72  |
| Qy | 661  | GATTAATGTTGATTTTGGGGCAACGCTGTGTCAGTGTGCTCAAGCATTTGCTTGGCTAA  | 720 |
| Db | 723  | GATTAATGTTGATTTTGGGGCAACGCTGTGTCAGTGTGCTCAAGCATTTGCTTGGCTAA  | 66  |
| Qy | 721  | GGCATGCTTGGTGTGATGATAGGCGGTTTGGCAATTTTCAAAAAATAGCAATTTTGTG   | 780 |
| Db | 663  | GGCATGCTTGGTGTGATGATAGGCGGTTTGGCAATTTTCAAAAAATAGCAATTTTGTG   | 60  |
| Qy | 781  | CAAAATATCCACATGCGATCGTGGCATCAAGATGATCAAGGAGAGATTTTAAATTTGT   | 840 |
| Db | 603  | CAAAATATATCCACATGCGATCGTGGCATCAAGATGATCAAGGAGAGATTTTAAATTTGT | 54  |
| Qy | 841  | TGCCCCCTTGACCTGCAATTTTGAATATATGCTCAAGCTGAAACCTTGACAGCTCAAGAC | 900 |
| Db | 543  | TGCCCCCTTGACCTGCAATTTTGAATATATGCTCAAGCTGAAACCTTGACAGCTCAAGAC | 48  |

|    |      |  |      |
|----|------|--|------|
| Qy | 901  | CACCAAAATCCATATTTGGATACGTTAATPAATCAAGGACAGGGCTGGCAATATTAACGGC  | 960  |
| Db | 483  | CACCAAAATCCATATTTGGATCGTTAATPAATCAAGGACAGGGCTGGCAATATTAACGGC   | 424  |
| Qy | 961  | GACACCAACAATCATGCGCTGCATGTTTGGCATGCGCTAATATGTCTGAACAGTCT       | 1020 |
| Db | 423  | GACACCAACAATCATGCGCTGCATGTTTGGCATGCGCTAATATGTCTGAACAGTCT       | 364  |
| Qy | 1021 | TTTGGCAATTTGAGCGCTGATGCGGAGATTTGGGATTTTGGCTTGAGTTAATCTCG       | 1080 |
| Db | 363  | TTTGGCAATTTGAGCGCTGATGCGGAGATTTGGGATTTTGGCTTGAGTTAATCTCG       | 304  |
| Qy | 1081 | TGCTTTGAGTGTATCAATGAACAATTTGAACATCGCTATCACAGAAATGCTTGAGCTTT    | 1140 |
| Db | 303  | TGCTTTGAGTGTATCAATGAACAATTTGAACATCGCTATCACAGAAATGCTTGAGCTTT    | 244  |
| Qy | 1141 | GGCGGCGACAATACTCGGCGCTTTCTTGGGTTAATGCCGGGCTGATGATCTGATCGGC     | 1200 |
| Db | 243  | GGCGGCGACAATACTCGGCGCTTTCTTGGGTTAATGCCGGGCTGATGATCTGATCGGC     | 184  |
| Qy | 1201 | TGAGCTTAAAGGTTCACTGTGCCAAGCTGCCAAAATGAACATGAACGGCTGTGGGCACTTT  | 1260 |
| Db | 183  | TGAGCTTAAAGGTTCACTGTGCCAAGCTGCCAAAATGAACATGAACGGCTGTGGGCACTTT  | 124  |
| Qy | 1261 | ATCTGCCAAGGTAGGATTTGGGATTTACCATCGATGATCTAACTTTGTGACCGGTATTTGAC | 1320 |
| Db | 123  | ATCTGCCAAGGTAGGATTTGGGATTTACCATCGATGATCTAACTTTGTGACCGGTATTTGAC | 64   |
| Qy | 1321 | CAAAAATTTCAACACTGACAGTCCAGAGCTGGCCAAACCGACGACGGCATATTTTTTTTGA  | 1380 |
| Db | 63   | CAAAAATTTCAACACTGACAGTCCAGAGCTGGCCAAACCGACGACGGCATATTTTTTTTGA  | 4    |

RESULT 2  
 US-10-672-787-35  
 : Sequence 35. Application US/10672787  
 : Publication No. US20040067554A1  
 : GENERAL INFORMATION:  
 : APPLICANT: LAGACE, Robert, E.  
 : APPLICANT: PATTERSON, Chandra  
 : APPLICANT: BERG, Kim, L.  
 : TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
 : FILE REFERENCE: ELITRA. 02561  
 : CURRENT APPLICATION NUMBER: US/10/672,787  
 : CURRENT FILING DATE: 2003-09-26  
 : PRIOR APPLICATION NUMBER: 09/596,002  
 : PRIOR FILING DATE: 2000-06-16  
 : NUMBER OF SEQ ID NOS: 41  
 : SOFTWARE: PERL Program  
 : SEQ ID NO 35  
 : LENGTH: 96109  
 : TYPE: DNA  
 : ORGANISM: Moraxella catarrhalis  
 US-10-672-787-35

Edmund

|    | Query Match           | Similarity            | Score 1380                 | DB 8                | Length 96109 |
|----|-----------------------|-----------------------|----------------------------|---------------------|--------------|
|    | Best Local Similarity | 100.0%                | Pred. No. 0                |                     |              |
|    | Matches 1380          | Conservative 0        | Mismatches 0               | Indels 0            | Gaps 0       |
| QY | 1                     | TAAAGCGTGAACCATGTAA   | CAAACTTATCACCCTGCCTCATTTAA | TTTTTTAAATTGATC     | 60           |
| Db | 11357                 | TAAAGCGTGAACCATGTAA   | CAAACTTATCACCCTGCCTCATTTAA | TTTTTTAAATTGATC     | 11410        |
| QY | 61                    | CATGCTTGACAAACGAGGTGA | CAAAAGCCCTTAGCGGTCTTGA     | CTGGCTGCCAA         | 120          |
| Db | 11417                 | CATGCTTGACAAACGAGGTGA | CAAAAGCCCTTAGCGGTCTTGA     | CTGGCTGCCAA         | 11476        |
| QY | 121                   | TACTGCCGATCAAGTGTAC   | CAACTGTATTTTAAAGTAGC       | CAAAATGTGCATCACCTTT | 180          |
| Db | 11477                 | TACTGCCGATCAAGTGTAC   | CAACTGTATTTTAAAGTAGC       | CAAAATGTGCATCACCTTT | 11533        |
| QY | 181                   | GAGTAAATCTTGCTCATTTA  | TGCGGGCATCTTGACGAGATTAA    | TACACACTATCCACCA    | 240          |



```
Db 11537 GAGTAATCTTGTCTCAATTTATCGGGGATCTTGACCGATTAATATACATATCCACCA 11596
Qy 241 CTCGATTAACATATGCTGACAGCTCTACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 300
Db 11597 CTCGATTAACATATGCTGACAGCTCTACTAAATCTTGAACCTTTACCAAGCCCAACCAAT 11656
Qy 301 CAATGCGAGTGAATATGCTCAATTAACCGTCCAAAGCCCAATTCATTCGATGTTGA 360
Db 11657 CAATGCGAGTGAATATGCTCAATTAACCGTCCAAAGCCCAATTCATTCGATGTTGA 11716
Qy 361 GCGGATGTTTGAACCTTTGAATCAATTAATATGCTTCATCAATATCAATATATA 420
Db 11717 GCGGATGTTTGAACCTTTGAATCAATTAATATGCTTCATCAATATCAATATATA 11776
Qy 421 TTTCACAGCGATGCGGTGAGCCCTTTGAAGTCTTAAGGTAATCAAGATGCGATTCATG 480
Db 11777 TTTCACAGCGATGCGGTGAGCCCTTTGAAGTCTTAAGGTAATCAAGATGCGATTCATG 11836
Qy 481 AAGATGGGAGCTGCTGCTAAGGCAAGTGAAGATTAAGGATTAAGGATTAAGCTTGGC 540
Db 11837 AAGATGGGAGCTGCTGCTAAGGCAAGTGAAGATTAAGGATTAAGGATTAAGCTTGGC 11896
Qy 541 TTGATTTTAAATGATGATGTTGAAGTCTTGAAGTCTTGAATATATATGATATACCATAC 600
Db 11897 TTGATTTTAAATGATGATGTTGAAGTCTTGAAGTCTTGAATATATATGATATACCATAC 11956
Qy 601 ATACCATCATCAATTAATTAATTAATTAATTCGCGTTTGGTGGCTAAGTATTTGTTGT 660
Db 11957 ATACCATCATCAATTAATTAATTAATTAATTCGCGTTTGGTGGCTAAGTATTTGTTGT 12016
Qy 661 GATTAATGATGATGTTGGGCAAGCTGTCAGTGTGTGAAGATGCTGTGCTAAAT 720
Db 12017 GATTAATGATGATGTTGGGCAAGCTGTCAGTGTGTGAAGATGCTGTGCTAAAT 12076
Qy 721 GGCATGCTTGTGATGATGATGCGGTTTGGCAATTTCAAAAATAGCAATTTTGTGC 780
Db 12077 GGCATGCTTGTGATGATGATGCGGTTTGGCAATTTTCAAAAATAGCAATTTTGTGC 12136
Qy 781 CAATTAATCCACATGCGATGCGATCAAGATGATCAAGCGAGATTAATTAATTTGT 840
Db 12137 CAATTAATCCACATGCGATGCGATCAAGATGATCAAGCGAGATTAATTAATTTGT 12196
Qy 841 TGCCCTTGAAGCTGCGCAATTTGAGATGCTCAAGTGAATCTTGAAGCTCAAGCAC 900
Db 12197 TGCCCTTGAAGCTGCGCAATTTGAGATGCTCAAGTGAATCTTGAAGCTCAAGCAC 12256
Qy 901 CACCAATCCATATTTTGAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
Db 12257 CACCAATCCATATTTTGAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12316
Qy 961 GACACCAACATATGCTGCGATGTTTGGCATCTGCGCTAATATGCTTAACAGTGT 1020
Db 12317 GACACCAACATATGCTGCGATGTTTGGCATCTGCGCTAATATGCTTAACAGTGT 12376
Qy 1021 TTGAGATTTGAAGCTGCTGATGCGGATGATGTTGGTGTGCTGATGATGATGCTG 1080
Db 12377 TTGAGATTTGAAGCTGCTGATGCGGATGATGTTGGTGTGCTGATGATGATGCTG 12436
Qy 1081 TGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 12437 TGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12496
Qy 1141 GGGGCGACATATCTGGGCGCTTCTGGGTTAATGCGGGGCTGATGATGATGATGATGATG 1200
Db 12497 GGGGCGACATATCTGGGCGCTTCTGGGTTAATGCGGGGCTGATGATGATGATGATGATG 12556
Qy 1201 TGAGCTTAAGATGCTGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCTGCTGCTGCT 1260
Db 12557 TGAGCTTAAGATGCTGCTGCAAGCTGCAAGCTGCAAGCTGCTGCTGCTGCTGCTGCT 12616
Qy 1261 ATCTGCCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
```

```
Db 12617 ATCTGCCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12676
Qy 1321 CAAAAAATTCACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 1380
Db 12677 CAAAAAATTCACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCTGACAGCT 12736

RESULT 3
US-10-282-122A-8940/c
; Sequence 8940, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8940
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-8940

Query Match 16.7%; Score 231; DB 8; Length 1344;
Best Local Similarity 52.4%; Pred. No. 4.1e-53;
Matches 721; Conservative 0; Mismatches 590; Indels 66; Gaps 7;
```

```
QY 182 AGTAAATCTTGTCAATTAATGAGGAGCATCTTGAACGATTAATAACACACTATCCACCAC 241
Db 1169 TGAATGGCTTGTGTGATGACCGGGGATCTTCAACATCAATACCAACATTTGGCATAT 1110
QY 242 TGCATTAACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAATC 301
Db 1109 TTTTCGTATGAAAGACGTAAGGAGAAATCTTGGCCCTTGGCTCCACCTAAATA 1050
QY 302 AATGCCA-----GTGATGATATGCCATTAACCGTGCACAGCCCATCAATTTGGCGCATG 355
Db 1049 AGTGCAACCTTACCTTTTTCATCTTAATGACGACACTTAAGCCATCAATTTGGCAAGT 990
QY 356 GTTGAAGCCGATGTTTGTACCTTTTGAATCAATTAATATGCTTATCAATATCATCA 415
Db 989 GTGACACCAACATTTGTACCTTTTGAATCAATTAATTAAGGACATACCCACAGTTTCA 930
QY 416 ATATATTACAGAGATGGGTAGCCCTTTGAAGTCTTAAAGGATATACAGATGGCATCC 475
Db 929 ACCTACTCACAGGATGCTCTTAATCTTTAAATGTTTAAAGTTTCACACATTTGACTCC 870
QY 476 ATGGGAAGATTGGAGCTGTGCTTAAGGCAAGTGCAGATGATAGGATGAGTATGTC 535
Db 869 ATAGTAACCAATTCCTCCCTCCCAATGCTAAACGCTAAAGCATTTGGCTACGTTGTC 810
QY 536 TTGCTTTGATTTTAAATGATGATGATTAAGATTTGTTTGTATGATATATATATGCC 595
Db 809 ATACCTTGAATATATTAATCCGAGCTTTTATTAACCGCTGTAAACACAGCAAGCCAA 750
QY 596 ATACCATCACATCATTAATTAATTAATTAATTCGCGCTTGTGCTTAACTAGCTATTT 655
Db 749 AGCGTACCGGCGATCTCTTAATAACGCAATATGAT-----TAAATCTGTGTCGT 698
QY 656 GTTGTATTAATGATGATTTGGGGCAACGCTTGTCACTGTCGTCAGCATGTCCTGGCT 715
Db 697 TTAAACCAAGCTTTGATTTGGTGTCTATCTGGAACAAGTGGCGGCTTA-----647
QY 716 AAATGGCATCGTCTTGTGATGATGATGCGGCTTTGGCAATTTTCAAAATATCGCAATTT 775
Db 646 ----AAGCATCGTCTGGTTAAATATACAACTTTTGTAGCGCTTGAATAATACATGTTTT 591
QY 776 TGTGCCAATTAATTCACATGCTCGTGGCGATCAAGATGATCAGGGAGAGTTTAA 835
Db 590 GCTTGTGATACCCAGCATATTTCCATGACGGTCTTAAGTGTCTTACTCATATTTAGA 531
QY 836 ATTTGTCCTTGTGAGTGCCTTAATTTGAATATGCTCAAGTAAATTTGACAGCTCA 895
Db 530 ACCACTGCTACCTCAGCGTTTAAAGTGGTCTTCTACTGAAAGCTTGATATCTCA 471
QY 896 AGCAGCAACCAATTCATATTTTGTGATGCTTAATTAATCAAGTGCAGGCGTGCATATTA 955
Db 470 AGGACTATGATGCTC---TGTGATCTTTAAGTAAATCTAAACGCTGTGACCAAGATTTG 414
QY 956 CCGCCGACACCAACATCAATGCTGCTGATGTTTGGCATCTCGCTACTAATATGCTTACA 1015
Db 413 CCGCCTTACTGCACTTTTCTTACCTGCACTCTTACGCAATTAAGCAATTAAGTGTACA 354
QY 1016 GTCTTTTGGCATTTGAGCTGTGATGATGAGTATGATGATGATGATGATGATGATGATG 1075
Db 353 GTACTTTTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
QY 1076 TCTGTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1135
Db 303 -----ACGACGCAATTAATTTGATATCACCCACACAGAAATATCTTTA 261
QY 1136 GCTTTGGCGGCAATATCTCGCGTCTTGGGTTATGCGGGGGCTGATGATGATGATGATG 1195
Db 260 GCAATAGCTGCTGAAATTTCCGTAATTTGTGTGCAAGGCTTGGGCTTAAATATATCTT 201
QY 1196 TCGGCTGAGCTTAAAGTACTGCTCAAGCTGCAAAATGACATGATGATGATGATGATG 1255
Db 200 TCTGCTTTGATATATATTTCTGATCAAGCTGACCAAAATCTGTTTAAAGCCGCGGA 141
QY 1256 AGTTTATCTGCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1315
```

```
Db 140 ATCTGATCGTGTCCGGAGGTGTGGGAGGAGTCCGTTACACCACTTGTAGCTTGT 81
QY 1316 TTGACCAAAAATTTACAGCTGACAGTCCAGACTGCCCAACGACGACGATAT 1372
Db 80 TCATGCAAGAAATTTTACAGCAAAACCTGATATTTCCAAAGCTGCTACACTTTT 24

RESULT 4
US-09-815-242-7090/C
? Sequence 7090, Application US/09815242
? Patent No. US20020061569A1
? GENERAL INFORMATION:
? APPLICANT: Haselbeck, Robert
? APPLICANT: Ohlsen, Karl L.
? APPLICANT: Zyskind, Judith W.
? APPLICANT: Wall, Daniel
? APPLICANT: Trawick, John D.
? APPLICANT: Carr, Grant J.
? APPLICANT: Yamamoto, Robert T.
? APPLICANT: Xu, H. Howard
? TITLE OF INVENTION: Identification of Essential Genes in
? TITLE OF INVENTION: Prokaryotes
? FILE REFERENCE: ELITRA.01A
? CURRENT APPLICATION NUMBER: US/09/815,242
? CURRENT FILING DATE: 2001-03-21
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? NUMBER OF SEQ. ID NOS: 1410
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ. ID NO 7090
? LENGTH: 1314
? TYPE: DNA
? ORGANISM: Haemophilus influenzae
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(1314)
? US-09-815-242-7090

Query Match 11.6%; Score 160.4; DB 3; Length 1314;
Best Local Similarity 49.6%; Pred. No. 2,1e-33;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

QY 1 TAAAGCGTGAACCATGTAAACAATTAATCACTGCTCATTTAAATTTTAAATGATC 60
Db 1308 TAAACATTTGAGTAAATGCGTAATTTCTTCCGCGCTTTTCAAAAGAACCAATGATC 1249
QY 61 CATGCTAGCACAAGCGGTGACAAAGACAGCCCTTAGCGGTCTTGACTGCTGCCAA 120
Db 1248 GAACTTTGCAAGCAGCGGACCAATTAATCA-----TATCTCCGCTTGGCA 1201
QY 121 TACTGCCGATCAAGTATACCAACCTGATATATTTAATGCAAAATGTCATCACCTT 180
Db 1200 TGTGGCGGTAAATTTCTATGCTTTTCCATTTGATGCAACA-----ATACTTTG 1147
QY 181 GAGTAAATCTTGTCTAATTAATCGGGGATCTTGACCGATTAATTAACACTATCACCA 240
Db 1146 CGATGAATAATTTTGAAGCAGCGCACCATCTGACCAAAACATTAATTAATGTTGG 1087
QY 241 CTGCAATTAATATGCTGACAGCTCAATTAATTTGAACCTTTTACCAAGCCCAACCAAT 300
```

1086 TTGATTAATTAATTCAGCTAATTCGAAAAATCAGCCCTTTTCGCTCCGCTAGCAA 1027  
QY 301 CAATGCCAGATGATATGCTTAACCGTCCAGGCCATTCATTCCTGATGTTGA 360  
DB 1026 CAATGCAATTTAACCTCAATATA-----AAGCCAGCCCATGACACCACTGTA 976  
QY 361 GCGGATGTTGATCTTTTGAATCAATTAATATGCTCTGATCATCATCATATA 420  
DB 975 CCCACATTTGTTGCTTTAGATCATTAACCAACGATGCAATGATGATGATGATA 916  
QY 421 TTCAACGCGATGCGGTAAGCTTTGAAGCTTAAGGATCAAGATGCGATCATGG 480  
DB 915 TTGAAAAAGATGATCTAACCTTTGAATGACGAAGCGGTAAGGATGATTAAT 856  
QY 481 AAGATGCGAGCTGCTGCTAAGGCAATGCAAGATGATGATGATGATGATGATG 540  
DB 855 AATACCTATGATGCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796  
QY 541 TTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 795 AACCAATGATGCTTTCAACAGGTAATATCACTTCACTTTCACTTAATATGCTT 736  
QY 601 ATCAACATCAATCAATTAATTAATTAATGTC-CGTTTGCTGCTAAGTAACTATTTG 659  
DB 735 GCGATTTTCAGTTTGTAGCCAAATATCGCATATTTCCGCAAAAGAAAGGATGTT 676  
QY 660 TGATATGCTGATTTGGGGAACGCTTGTCAATGCTGCTGCTGCTGCTGCTGCT 719  
DB 675 CGCTGATTTTCTGTTTCCCAAAAGTCAAGC----- 644  
QY 720 TGGCATGCTTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 779  
DB 643 ---TATCTTCAATTTTCAACACACTTACTTGAATTAATTAATGCTTAATTTGCTT 587  
QY 780 CCAATTAATCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839  
DB 586 GGGCATATCTTCAATTCATATAGCATATGATGATGATGATGATGATGATGATG 527  
QY 840 TTGCGCTTGAAGCTGCAATTTGATGATGATGATGATGATGATGATGATGATG 899  
DB 526 TCGCTGCGAGCTTTTAAGCTATAGTGTCTCAACCTGAAAGCTGATGTA 467  
QY 900 CCAACCAATCCATTTTGGATGCTTAATTAATTAATTAATTAATTAATTAATTA 959  
DB 466 CATTAATGATCAATCTTCAATC---AACATGCAAGGAGGAAATCCATATTTCCG 410  
QY 960 CGACCAACCAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
DB 409 CCAATACCAATTTCAACACGAGCTTTCGCTTTCATTAATTAATTAATTAATTA 350  
QY 1020 TTTTGGATTTGAGCTGCTGATGCGGATGATGATGATGATGATGATGATGAT 1079  
DB 349 TTTTACATTTGAACCTGTAATCCCACAATGCTTTGCTGCGCGGCA----- 298  
QY 1080 GTGCTTGAAGTATCAATGAACATGATGATGATGATGATGATGATGATGATG 1139  
DB 297 -----GATTAATTCATATGCGATTAATCTTCCATCCGCTTAA 257  
QY 1140 TGGGCGGCAATATCTGCGGCTTCTGCTTAATGCGGCTGATGATGATGATGAT 1199  
DB 256 GTGCGGTTTGAATTTCTGCTTTTTCACGCAAGCCCTGCTTAATCAATCATTCG 197  
QY 1200 CTGAGCTTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259  
DB 196 TTTCAATTAACATTTCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 137  
QY 1260 TATTCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1319  
DB 136 TATTAATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATG 77  
QY 1320 CCAAAAAATTCACGCTGACAGTCCAGAGCTGCCAAACGACGACGCTATTTT 1377  
DB 76 AGAGATATTCACACAGAAAGACCTGTTTTCAGAGCCGATGATGATATATTTT 19

RESULT 5  
US-10-282-122A-22236/c  
; Sequence 22236, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: EPIPA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: Patencin version 3.1  
; SEQ ID NO 22236  
; LENGTH: 1314  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
US-10-282-122A-22236  
Query Match 11.6%; Score 160.4; DB 8; Length 1314;  
Best Local Similarity 49.6%; Pred. No. 2.1e-33;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
QY 1 TAAAGCTGAACCATGTTAACAACCTATACCTGCTCATTTAATTTTAATGATC 60  
DB 1308 TAAACATGAGCTTAATGCGTAATTTCTGCGGCTTTTCAAAAGAACATGAT 1249  
QY 61 CATGCTGACACAGAGGTGACAAAGACACAGCCCTAGCGGCTTTGACTGCTGCA 120  
DB 1248 GAGACTTGCAAGAGCGACGACATTAATCA-----TATCTCCGCTTGCA 1201  
QY 121 TACTGCCGATCAAGTGTACCACTGATATATTTAGATGCAAAATGCTACCTT 180  
DB 1200 TGTGGGCGTAATAAATTTCTATCCCTTGTTCATGTAATGCAACA-----ATACTT 1147  
QY 181 GATTAATCTTCTCAATTTATGCGGATGATGATGATGATGATGATGATGATGAT 240  
DB 1146 CGATGAATAATTTTGCAAGCGCCACCATCTGACCAAAACATTAATTAATGCTG 1087  
QY 241 CTGATTAACATATGCTGACAGCTCACTAAATCTTGAACCTTTACAAAGCCCAAAAT 300

```

Db      1086 TTGATTAATTAATTCAGCTAATTCGAAAATCAGCCCTTTCCGCTCCGCTGACAA 1027
Qy      301 CAATGCCAGTGATGATATGCCATTAACCGTCCAGGCCCATCAATGCTGCAATGGTTGA 360
Db      1026 CAATGCAATTTTACCTTCATATA-----AAGCCAGCCCAATGACCACTGACT 976
Qy      361 GCGAGTGTGTACCTTTTAATCATTAATAATGCTTGCATCAATATCAATCAATATA 420
Db      975 CCCCACTTTGTTGCTTTAGATCATTAATCCAAAGATCCATTAAGCTTGAATGACTTA 916
Qy      421 TTCAACGATGAGGTGAGCCCTTTGAAAGTCTTAAGGATCAAGATGACATCCATGGG 480
Db      915 TTGAAAACGATGATCTAACCTTTGAAATGACGAAGCGGTACGATGATGAATCTAAAT 856
Qy      481 AAGATTGGACGCTGTGCTTAAGGCAATGACGATAGGCAATGATGATGATGCTTGGC 540
Db      855 AATACCTATAGCTTGTCCAATGCTGTGCTGCCAAATGTTCTATATATATATAGCGACC 796
Qy      541 TTGATTTTAAATTTGATGATGATGATTAAGTTGTTTGTGATGATATATAGCATACC 600
Db      795 AACCAATGATGCTTTCACAGGATAAATCACTTCACTTTACATTAATATGCTT 736
Qy      601 ATCAACATCATCAATTAATTAATTAATCTGC-CGTTTGGGTGCTTAAGTATGTTG 659
Db      735 GCCATTTTCACTTTTATGCAATATATCCGACTATTTTCGCAAAAGAAAGGATGTTT 676
Qy      660 TGAATATGATGATTTGGGGCAACGCTTGTCACTGTGTCACATGCTTGGCTAAT 719
Db      675 CGCTGTATTTTCTTTTCCCAAAAGTCAGC----- 644
Qy      720 TGGCATGCTTGGTATGATGATGCGGTTGGCAATTTTCAAAATAGCAATTTTGTG 779
Db      643 ---TATTTATTTGTTCAACACACTTCTTAGCATTAATGATTAATGCTAATTTTCTT 587
Qy      780 CCAATTAATCCACCATGCAATCGTCCGATCAAGATGATGACGAGAGATTTAAATG 839
Db      586 GGGGATATCTTTCAATCCATATAGGATCATATATCTTTCAGTCAAGTCAAGACAG 527
Qy      840 TTGCCCCCTTGAAGTCCAAATTTGATATGCTCTAAGCTGAAAATTGACAGCTCAAGCA 899
Db      526 TCCTGTCGAGCTTTTAAGCTATAGTGTCTCAAGCTGAAAATTGAAAAGCTTGA 467
Qy      900 CCAACCAATCCATTTTGTGATGCTTATATTAATCAAGTCAAGGCGTCCATATTACCG 959
Db      466 CATTAATTCATCTTTCATTC---AACATGCAAAAGCGGGAATCCATATTTTCCG 410
Qy      960 CGACACCAACATATGCTGATGTTTGTGCAATCTGCTTATATGCTGATACAGTGC 1019
Db      409 CCAATACCACTTTCAACCAAGACCTTTCCGCAATTAATAAAGTATGCGTAC 350
Qy      1020 TTTTGGCAATTTGACCTGTGATGCGGATGATGCTGTGTTGTTGCTTGAAGTATC 1079
Db      349 TTTTACCATTTGAACCTGTATATCCCAATTTGCTTGTGCTGCGCGGCA----- 298
Qy      1080 GTGCTTTGAGTATCAATGAACAATGAACATGCTGATCAGAGATGCTTGAGCTT 1139
Db      297 -----GATTAATTCATATGCGGATTAATCTTCCATCCCTTTAA 257
Qy      1140 TGGCGGCAATATCTGCGCTTCTTGGGTAAATGCGGGCTGATGATGATGATCGG 1199
Db      256 GTGCGGTTTGAATTTCTGCTGTTTAAACCGCAAGCCCTGGGCTAATATCATATCGC 197
Qy      1200 CTGAGCTTAAGATTTCACTGTCCAAAGTCCAAATATGACATGAAGCGCTGTGGCGATT 1259
Db      196 TTTCAAGTAAACATTTCTGATTTAACTAACAGATGAAGGAGATATTTTGAAGAAATT 137
Qy      1260 TATCTGCAAGTATGATTTGGGATTTACCATGATGATGATCACTTTGACCGTATGTA 1319
Db      136 TATTAATAACGATGATTTTTCGAGATTAATCAACGAATTTAGCTGTGGGATA 77
Qy      1320 CCAAAAATTCAGCTGACAGTCCAAGACTGCCAAACGACGACGGCATATTTT 1377

```

```

Db      76 AGAGATATATCCACACAGAAGAAGCTGTTTGGCAAGCCCGATGATGATATATTTT 19
RESULT 6
US-10-958-216-400/c
; Sequence 400, Application US/10958216
; Publication No. US2005018138A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ALBED
; APPLICANT: DHARMSI, AKIL
; APPLICANT: VEDADI, MASOUD
; APPLICANT: ALAM, MUHAMMAD ZAHOR
; APPLICANT: AROSMITH, CHERYL
; APPLICANT: AWEY, DONALD E.
; APPLICANT: BEATTIE, BRYAN
; APPLICANT: BUZADZID, KRISTINA
; APPLICANT: CANADEN, VERONICA
; APPLICANT: DOMAGALA, MEGAN
; APPLICANT: HOUSTON, SIMON
; APPLICANT: KANAGARAJAH, DHUSHY
; APPLICANT: LI, QIN
; APPLICANT: MANSOURY, KAMRAN
; APPLICANT: McDONALD, MERRY-LYNN
; APPLICANT: NETHERY, KATHLEEN
; APPLICANT: NG, IVY
; APPLICANT: OUYANG, HUI
; APPLICANT: PINDER, BENJAMIN
; APPLICANT: RICHARDS, DAWN
; APPLICANT: TAI, MATTHEW
; APPLICANT: THALAKADA, ROSANNE
; APPLICANT: VALLEE, FRANCOIS
; APPLICANT: VIRAG, CRISTINA
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM BACTERIA
; FILE REFERENCE: IPT-205.01
; CURRENT APPLICATION NUMBER: US/10/958, 216
; CURRENT FILING DATE: 2004-10-04
; PRIOR APPLICATION NUMBER: PCT/CA03/00462
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/369, 511
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 60/385, 089
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/385, 751
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/386, 553
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386, 577
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386, 367
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386, 566
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386, 390
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386, 601
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1132
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 400
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-958-216-400

Query Match      11.6%; Score 160.4; DB 10; Length 1314;
Best Local Similarity 49.6%; Pred. No. 2.1e-33;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy      1 TAAAGCTGAACATGTTAACAACCTATATCACTGCTCATTAATTTTAATGATC 60
Db      1308 TAAACATGAGCTTAATGCTTAATTTCTTGGCGGCTTTTCAAAAGACCAATGATC 1249

```

QY 61 CATCTAGACAGACAGAGTGCAGAAAGACAGCCCTAGCGGTGCTTGAATGCTGCGCAA 120  
DB 1248 GAGCTTGCACAGACAGAGCCAGCAATTAATACCA-----TATCTCGCTTTGGCA 1201  
QY 121 TACTGCCGATCAAGTGTACCACTGATATATTTAGATGCCAATGTGCATCACTTT 180  
DB 1200 TGTGGGCGTAAATAATCTATCGCTGTCTGATGATGACAA-----ATACTTTG 1147  
QY 181 GAGTAATCTTGTCTCAATTAATCGGGGATCTTGACCGATTAATACACATCTACCCA 240  
DB 1146 CGATGAAAAATTTGCAAGCAGCGACCATCTGCACAAACAAATTAACAAATATGTGTG 1087  
QY 241 CTGCATTAACATATCTGACAGCTCACTAAATTTTGAACCTTACCAAGCCCAACAAAT 300  
DB 1086 TTGATTAATTAATTCAGCTAATCTGAAAAATCAGCCCTTTTCGCTCTCGCCTAGCA 1027  
QY 301 CAATGCGAGTATGATATGCCATTAACCGTGCAGGCCATCAATTCCTGCATGTTGA 360  
DB 1026 CAATGCAATTTACCTCAATATA-----AAGCCAGCCCAATGACAGCACTGACT 976  
QY 361 GCGGATTTTGTACTTTTGAATGATTAATAATGCTTGTATCAATATCATATA 420  
DB 975 CCCCAATTTGTGCTTGAAGTATTAATCAACGAATGCAATTAATCTGATGACTTAA 916  
QY 421 TTCAACGAGTACGCTGACCTTTGAAAGTCTTAAGGATCAAGCATGCGATCGATGG 480  
DB 915 TTGAAAACGATGATCTAACCTTTGAAATGACGAAGTCCGTAAGATTAATTAAT 856  
QY 481 AAGATTGCACTGTGCTTAAGGCAAGTGCAGATTAAGGCAATGATGATGCTTGGC 540  
DB 855 AATCTATAGCTGTGCTCAATGCTGTGCGCAAAATGTTCAATTAATATGCGACG 796  
QY 541 TTGATTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
DB 795 AACCAATGATGCTTCTCAAGGTAATTAATCACTTAATTAATTAATTAATTAAT 736  
QY 601 ATCAACATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 659  
DB 735 GCAATTTTGAATTTTGAACGATTAATGCGCATTTTTCGCAAAAGCAAGATGAT 676  
QY 660 TGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719  
DB 675 CGCTGATTTGCTTTTCCCAAAAGTCAAGC----- 644  
QY 720 TGGGATGCTTGTATATATATGCGGTTTGGCAATTTTCAAAATACGAATTTTGTG 779  
DB 643 ---TATCTATGTTTCAACACACTTACCTTAATTAATTAATTAATTAATTAAT 587  
QY 780 CCAATATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839  
DB 586 GGGGATTAATTTCTAATTCATATAGCATTCATTAATTTCTGCTACGTTCAAGCAG 527  
QY 840 TTGCGCTTGAAGTGCAGAAATTTGATATGCTCAAGCTGAAAATTTGACGCTCAAG 899  
DB 526 TCGCTGCGGAGCTTTTAAGCTATAAGTGTCTCAAGCTGAAAATTTGAAAAGCTT 467  
QY 900 CCACCAATGCTAATTTTGAATGCTTAATTAATTAATTAATTAATTAATTAATTA 959  
DB 466 CATTAATTTCAATTTCAATTC---AACATGACAAAGCGGGAATCCCAATTTTTCG 410  
QY 960 CGACCAACCAATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1019  
DB 409 CCAATCAATTTTCAACGACGAGCTTTCGCAATTTTCAATTAATTAATTAATTAAT 350  
QY 1020 TTTTGAATTTGAAGCTGTGATGCGATGATGATGATGATGATGATGATGATGATG 1079  
DB 349 TTTTACATTTGAACCTGATATCCCAATTTGCTTGTGCTGCGCGCA----- 298  
QY 1080 GTGCTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139  
DB 297 -----GATTAATTTCAATTTGCGGATGATGATGATGATGATGATGATGATG 257  
QY 1140 TGGCGGAGCAATATCTGCGGCTTCTTGATGATGATGATGATGATGATGATGATG 1199

DB 256 GTGCGGTTGAATTTCTGATGATTTTACCGCAAGCCCTGGGCTAATTAATCAATATCGC 197  
QY 1200 CTGAGCTTAAGAGTTCACTGTCCAGCTGCCAAATATGACATGACCGCTGTGGCAGTT 1259  
DB 196 TTTCAGTAACTCAATCTCTGATTTTAACTACAGTATGACAAAGGATATTTTGAAGATT 137  
QY 1260 TATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319  
DB 136 TATCAATATCAAGTATGATTTTTCGATATCAATACAGAAATTAATGCTGTGGGATA 77  
QY 1320 CCAAAAAATTTCAAGCTGACAGTCCAGAGCTGCCAAACGAGACGCGATATTTT 1377  
DB 76 AGATATATCCACAGAAAGACGCTTTTGCAGACCGGATGATGATATTTT 19  
  
RESULT 7  
US-11-194-246-220/c  
; Sequence 220, Application US/11194246  
; Publication No. US20050272089A1  
; GENERAL INFORMATION:  
; APPLICANT: Mott, John  
; APPLICANT: Trepod, Catherine  
; APPLICANT: Arvidson, Scaffan  
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METH.  
; FILE REFERENCE: 00592.US1 (M&P 268, 05920101)  
; CURRENT APPLICATION NUMBER: US/11/194,246  
; CURRENT FILING DATE: 2005-08-01  
; PRIOR APPLICATION NUMBER: US/10/274,586  
; PRIOR FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: US 60/345,438  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 621  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 220  
; LENGTH: 2925  
; TYPE: DNA  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: Sequence of the mureD coding sequence and flanking regions.  
US-11-194-246-220  
  
Query Match 11.6%; Score 160.4; DB 15; Length 2925;  
Best Local Similarity 49.6%; Pred. No. 3.4e-33;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;  
  
QY 1 TAAAGCTGAACCATGTTAAACAATTAATCACTGCTCATATTAATTTTAATGATC 60  
DB 2240 TAAACATGAGCTAATGCTAATTTCTTGGCGGCTTTTCAAAAGCAAACTGATC 2181  
QY 61 CATCTAGACAGACAGTGCAGAAAGACAGCCCTAGCGGTGCTTGAATGCTGCGCAA 120  
DB 2180 GAGACTTGCACAGACAGCGCCACATTAATACCA-----TATCTCGCTTTGCAA 2133  
QY 121 TACTGCCGATCAAGTATGACCACTGATATTTAGATGCCAATGTGCATCACTTT 180  
DB 2132 TGTGGGCGTAAATAATCTATCGCTGTCTGATGATGATGATGATGATGATGATG 2079  
QY 181 GAGTAATCTTGTCTCAATTAATCGGGGATCTTGACCGATTAATTAATCACTATCCA 240  
DB 2078 CGATGAAAAATTTGCAAGCAGCGACCATCTGACCAAAACAAATTAATGTGTG 2019  
QY 241 CTGCATTAACATATGCTGACAGCTCACTAAATTTTGAACCTTACCAAGCCCAACAA 300  
DB 2018 TTGATTAATTAATTAATGCTAATCTGAAAAATGAGCCCTTTTCGCTCTCGCTG 1959  
QY 301 CAATGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
DB 1958 CAATGCAATTTACCTCAATATA-----AAGCCAGCCCAATGACGAACTGACT 1908  
QY 361 GCGGATTTTGAATCTTTGAATCAATTAATTAATGCTTGTCTGATCAATATCATATA 420





|   |      |  |      |
|---|------|--|------|
| Db  | 8400 | GGCGATATATCTTTAAATCCATATAGGATCCATATGATCTTACGTACAGCTTCAGACAG    | 8341 |
| Qy  | 840  | TTGCCCTTGAAGTCCGCAATTTGAGATATGCTCAAGCTGAAATCTGACAGCTCAAGCA     | 899  |
| Db  | 8340 | TCGCTCGCGAGCTTTTAAAGCTATTAAGTGTCTCAAGCTGAAACTGAAAGCTCTAAGTA    | 8281 |
| Qy  | 900  | CCACCAATCCATATTTTGGATCGTTAATAATTCAGTGAAGCGCGTCCAAATTTTACCGC    | 959  |
| Db  | 8280 | CATAAAGTTCAACAATCTTCAATTC---AACAATGACAAAGGGGAATCCCAATATTTTCCCG | 8224 |
| Qy  | 960  | CGACACCAACATCATGCTGATGCTGATGTTTGGCATCTGCGCTTAATATGTCGTAACAGTGC | 1019 |
| Db  | 8223 | CCATACCAACTTTACACACGAGAGCTTTCCGCATTTTATTAACCTAAAGTATGTTACGGTAC | 8164 |
| Qy  | 1020 | TTTTGGCATTTGAGCCCTGTGATGATGCGATGATGTTGGTGTGTTTGTGCTTGATTAATCTC | 1079 |
| Db  | 8163 | TTTTACCATTTGAACTGTATATCCCAATTTGGCTTTGTGCGTGGCGGCA-----         | 8112 |
| Qy  | 1080 | GTCGCTTGAGTGTATCAATGAACATTCGCTGATCAAGAAATGCTTGAGCTT            | 1139 |
| Db  | 8111 | -----GAAATATTCATATATCGCGATTAATCTTCCATCCCGCTTTTAA               | 8071 |
| Qy  | 1140 | TGGGGGGGACATATCTGGGGTCTTGGGTTAATGCCGGGCGATGATGATCTGATCGG       | 1199 |
| Db  | 8070 | GTCGGGGTTGATATTTCTGGTGTGTTTTCGCGACAGCCCTGGGCTAAATACATATATCTGC  | 8011 |
| Qy  | 1200 | CTGAGCTTAAAGATTCAGTCTGTCGACAGTCCGCAAAATGACATGAACGCTGTGGGCAATT  | 1259 |
| Db  | 8010 | TTTCAAGTAAACCATTTCTGATTTTAATACAGATATGAAGAGGATATTTTGAAGAAATT    | 7951 |
| Qy  | 1260 | TATCTGCCAAGTATGATGGGATTTACATGATGATGCTTAACCTTTGTGACCGTGAATTGA   | 1319 |
| Db  | 7950 | TATCAATATCACAGTGAATTTTTCGAGTATCAATACACAGAAATATTAAGCCTGTGGGATA  | 7891 |
| Qy  | 1320 | CCAAAAATTCACAGCTGACATGCCAGAGTCCGCAACCGCAGAGCAGCATATTTT         | 1377 |
| Db  | 7890 | AGAGATATATCCACACAGAAAGACCTGTGTTTGGCCAGCCGATGATGATATATTTT       | 7833 |
| RESULT 9  |      |  |      |
| US-10-329-670-1/c   |      |  |      |
| Sequence 1, Application US/10329670   |      |  |      |
| Publication No. US20040018503A1   |      |  |      |
| GENERAL INFORMATION:  |      |  |      |
| APPLICANT: Fleischmann et al.   |      |  |      |
| TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag |      |  |      |
| TITLE OR INVENTION: Theoreof, and Uses thereof  |      |  |      |
| FILE REFERENCE: PBI86P1   |      |  |      |
| CURRENT APPLICATION NUMBER: US/10/329, 670  |      |  |      |
| PRIORITY FILING DATE: 2002-12-24  |      |  |      |
| PRIOR APPLICATION NUMBER: US 09/643, 990  |      |  |      |
| PRIOR FILING DATE: 2000-08-23   |      |  |      |
| PRIOR APPLICATION NUMBER: US 08/487, 429  |      |  |      |
| PRIOR FILING DATE: 1995-06-07   |      |  |      |
| PRIOR APPLICATION NUMBER: US 08/426, 787  |      |  |      |
| PRIOR FILING DATE: 1995-04-21   |      |  |      |
| NUMBER OF SEQ ID NOS: 1   |      |  |      |
| SOFTWARE: PatentIn version 3.1  |      |  |      |
| SEQ ID NO 1   |      |  |      |
| LENGTH: 1830121   |      |  |      |
| TYPE: DNA   |      |  |      |
| ORGANISM: Haemophilus influenzae  |      |  |      |
| FEATURE:  |      |  |      |
| NAME/KEY: misc_feature  |      |  |      |
| LOCATION: (4747)..(4747)  |      |  |      |
| OTHER INFORMATION: n equals a, t, g or c  |      |  |      |
| FEATURE:  |      |  |      |
| NAME/KEY: misc_feature  |      |  |      |
| LOCATION: (9921)..(9921)  |      |  |      |
| OTHER INFORMATION: n equals a, t, g or c  |      |  |      |
| FEATURE:  |      |  |      |
| NAME/KEY: misc_feature  |      |  |      |
| LOCATION: (10150)..(10150)  |      |  |      |
| OTHER INFORMATION: n equals a, t, g or c  |      |  |      |



```
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (80024)..(80024)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (100091)..(100091)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (102656)..(102656)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (105121)..(105121)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (107248)..(107248)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (117136)..(117136)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (119750)..(119750)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (119924)..(119924)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (120038)..(120038)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (121344)..(121344)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (122167)..(122167)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (122336)..(122336)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (131340)..(131340)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (131360)..(131360)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (139910)..(139910)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (140398)..(140398)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (142750)..(142750)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (145058)..(145058)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
```

```
? NAME/KEY: misc_feature
? LOCATION: (145171)..(145171)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (145942)..(145942)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (147197)..(147197)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (150841)..(150841)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (152500)..(152500)
? OTHER INFORMATION: n equals a, c, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (152530)..(152530)
```

Query Match 11.6%; Score 160.4; DB 8; Length 1830121;  
Best Local Similarity 49.6%; Pred. No. 1.4e-31;  
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy 1 TAAAGGTGACAGCATTTAAACAACTATCCAGCTCATTTAAATTTTAAATGATC 60
Db 1204439 TAAACATGAGCTTAAATGCGTAAATCTTTCGCCGCTTTCAAAAGAACAACTGATC 1204380
Qy 61 CATGCTAGCACAAGCAGTGACAAAAGCAGCCCTAGCGGTGCTTGACTGCTGCCAA 120
Db 1204379 GAGACTTGACACAGAGGAGCAATATACCA-----TATCTCCGCTTGCAA 1204332
Qy 121 TACTGCCGATCAAGGTACCAACTGATATATTTAGATCCAAATGTCATCCTTT 180
Db 1204331 TGTGGGCGTAAAAATTCATCCCTGTTCCATGTATGACAA-----ATACCTTGG 1204278
Qy 181 GAGTAAATCTGCTCAATTAATGCGGCGATCTTACCGATTAATATACACTATCCACCA 240
Db 1204277 CGATGAAATTTTGCACAGCGCCACCATCTGCACCAACATTAATATATGTGTG 1204218
Qy 241 CTCATPAACATATGCTGACAGCTCACTAAATCTTGACCTTTACCAAGCCCAACCAAAAT 300
Db 1204217 TTGATTAATTAATTCAGCTAATTTCTGAAAATAGCCCTTTCCGCTCCGCTGCAA 1204158
Qy 301 CAATGCCAGTGAATGATGCAATAAACCGTCCAAAGCCCATCAATTGCTGCAATGTTGA 360
Db 1204157 CAATGCAATTTACCTCAATATA-----AAGCCCAAGCAATGCAACACTGTACT 1204107
Qy 361 GCCGATGTTTGTACCTTTTGAATCATTAATAATATGCTGCTTCAATATCAATATATA 420
Db 1204106 CCCCAATTTTGTGCTTTTGAAGTCAATTAATCCACGAATGCCATTAGCTGATGACTAA 1204047
Qy 421 TTCACAGCATGCGGTAGCCCTTTGAAAGCTTTAAGGATATCAAGCATGCGATCCATGGG 480
Db 1204046 TTGAAAACGATGATCTAACCTTTGAAATGACGAAGCGGTGCAATGAATCTAAAT 1203987
Qy 481 AAGATTGGCAGCTGTGCTTAAAGCAAGTCAATTAAGCATTTAGTAGGTTATGCTTGGC 540
Db 1203986 AATACCTATGCTGTGCGCAATCTGTGCTGCAAAATGTTATTAATTAATGCGGACC 1203927
Qy 541 TTTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 1203926 AACCAATGATGCTTCTTCAAGATTAATCACTTATCTTTTAAATTAATGCTT 1203867
Qy 601 ATCACCATCATCAATTAATTAATTAATCTGC-CGTTGTGCTGCTAGTAAGCTATTTGTTG 659
Db 1203866 GCCATTTTCAAGTTTAAAGCAATTAATCCGCAATATTTCCGCAAAAGAAACGATATGTT 1203807
Qy 660 TGATATGCTGATTTTGGGCAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
```

|                    |                    |        |        |        |      |
|--------------------|--------------------|--------|--------|--------|------|
|                    | OTHER INFORMATION: | n      | equals | a,t,c, | or g |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (9921)..(9921)     |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (10150)..(10150)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (29298)..(29298)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (36543)..(36543)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (36551)..(36551)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (36636)..(36636)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (40808)..(40810)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (44416)..(44416)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (44905)..(44905)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (44975)..(44975)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (45593)..(45593)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (45732)..(45732)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (47036)..(47036)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (51602)..(51602)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (51786)..(51786)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (51805)..(51805)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |
| FEATURE:           |                    |        |        |        |      |
| NAME/KEY:          | misc_feature       |        |        |        |      |
| LOCATION:          | (55369)..(55369)   |        |        |        |      |
| OTHER INFORMATION: | n                  | equals | a,t,c, | or g   |      |

```

1 FEATURE:
2 NAME/KEY: misc_feature
3 LOCATION: (65309) .. (65309)
4 OTHER INFORMATION: n equals a,t,c, or g
5
6 FEATURE:
7 NAME/KEY: misc_feature
8 LOCATION: (65313) .. (65113)
9 OTHER INFORMATION: n equals a,t,c, or g
10
11 FEATURE:
12 NAME/KEY: misc_feature
13 LOCATION: (80024) .. (80024)
14 OTHER INFORMATION: n equals a,t,c, or g
15
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: (10091) .. (10091)
19 OTHER INFORMATION: n equals a,t,c, or g
20
21 FEATURE:
22 NAME/KEY: misc_feature
23 LOCATION: (102696) .. (102696)
24 OTHER INFORMATION: n equals a,t,c, or g
25
26 FEATURE:
27 NAME/KEY: misc_feature
28 LOCATION: (105121) .. (105121)
29 OTHER INFORMATION: n equals a,t,c, or g
30
31 FEATURE:
32 NAME/KEY: misc_feature
33 LOCATION: (107248) .. (107248)
34 OTHER INFORMATION: n equals a,t,c, or g
35
36 FEATURE:
37 NAME/KEY: misc_feature
38 LOCATION: (117136) .. (117136)
39 OTHER INFORMATION: n equals a,t,c, or g
40
41 FEATURE:
42 NAME/KEY: misc_feature
43 LOCATION: (119750) .. (119750)
44 OTHER INFORMATION: n equals a,t,c, or g
45
46 FEATURE:
47 NAME/KEY: misc_feature
48 LOCATION: (119924) .. (119924)
49 OTHER INFORMATION: n equals a,t,c, or g
50
51 FEATURE:
52 NAME/KEY: misc_feature
53 LOCATION: (120038) .. (120038)
54 OTHER INFORMATION: n equals a,t,c, or g
55
56 FEATURE:
57 NAME/KEY: misc_feature
58 LOCATION: (122167) .. (122167)
59 OTHER INFORMATION: n equals a,t,c, or g
60
61 FEATURE:
62 NAME/KEY: misc_feature
63 LOCATION: (122336) .. (122336)
64 OTHER INFORMATION: n equals a,t,c, or g
65
66 FEATURE:
67 NAME/KEY: misc_feature
68 LOCATION: (131340) .. (131340)
69 OTHER INFORMATION: n equals a,t,c, or g
70
71 FEATURE:
72 NAME/KEY: misc_feature
73 LOCATION: (131360) .. (131360)
74 OTHER INFORMATION: n equals a,t,c, or g
75
76 FEATURE:
77 NAME/KEY: misc_feature
78 LOCATION: (139910) .. (139910)
79 OTHER INFORMATION: n equals a,t,c, or g
80
81 FEATURE:
82 NAME/KEY: misc_feature
83 LOCATION: (140398) .. (140398)
84 OTHER INFORMATION: n equals a,t,c, or g
85

```

```

1 NAME/KEY: misc_feature
2 LOCATION: (142750) ..(142750)
3 OTHER_INFORMATION: n equals a,t,c, or g
4 FEATURE:
5 NAME/KEY: misc_feature
6 LOCATION: (145058) ..(145058)
7 OTHER_INFORMATION: n equals a,t,c, or g
8 FEATURE:
9 NAME/KEY: misc_feature
10 LOCATION: (145171) ..(145171)
11 OTHER_INFORMATION: n equals a,t,c, or g
12 FEATURE:
13 NAME/KEY: misc_feature
14 LOCATION: (145942) ..(145942)
15 OTHER_INFORMATION: n equals a,t,c, or g
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: (147197) ..(147197)
19 OTHER_INFORMATION: n equals a,t,c, or g
20 FEATURE:
21 NAME/KEY: misc_feature
22 LOCATION: (150841) ..(150841)
23 OTHER_INFORMATION: n equals a,t,c, or g
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: (152500) ..(152500)
27 OTHER_INFORMATION: n equals a,t,c, or g
28 FEATURE:
29 NAME/KEY: misc_feature
30 LOCATION: (152530) ..(152530)

```

|                       |                 |                    |            |                 |
|-----------------------|-----------------|--------------------|------------|-----------------|
| Query Match           | 11.6%           | Score 160.4;       | DB 9,      | Length 1830121; |
| Best Local Similarity | 49.6%           | Pred. No. 1.4e-31; |            |                 |
| Matches 683;          | Conservative 0; | Mismatches 606;    | Indels 89; | Gaps 77;        |

|    |         |               |              |                          |                            |               |                       |                 |
|----|---------|---------------|--------------|--------------------------|----------------------------|---------------|-----------------------|-----------------|
| QY | 1       | TAAGCGTGAA    | CCATGTTAA    | CAAACTTAT                | CACCTGCTCAT                | TATATTTTAA    | TTGATC                | 60              |
| Db | 1204439 | TAACATTTGAGCT | AAATGCGTAAAT | CTTGCGCCGCGCTTTT         | CAAAAGCAAACTATC            |               | 1204380               |                 |
| QY | 61      | CATCTGAGCA    | CABA         | CAGGTGA                  | CAAA                       | AGCAGCCCTG    | AGCGGTGCTTGAC         | GTGCGCA 120     |
| Db | 1204379 | GAGACTTGCA    | CAGACGCGCA   | CAATTA                   | TACCA-----TATCTCGCTTGCA    |               | 1204332               |                 |
| QY | 121     | TACAGCCCGAT   | CAAGTGM      | CAACCTGAT                | TATATTTAGATG               | CCAAATGTGCAT  | CACCTTT 180           |                 |
| Db | 1204331 | TGTTGGGCGT    | AAAAATCT     | ATCGTGTTCAT              | TGTATCGAACAA-----ATAACTTTG |               | 1204278               |                 |
| QY | 181     | GAGTAAATCT    | TGCTCA       | TTATCGGGG                | CACTTTGACGAT               | TAAATACACAT   | ATACACCA 240          |                 |
| Db | 1204277 | CGATGAAATTT   | TGCAAG       | CAGCGCACCAT              | CTCGACCAAA                 | CAATATATATG   | TGTGG 1204218         |                 |
| QY | 241     | CTGCATACAT    | ATGCTGA      | CAGCTCAT                 | TAATCTTTGACCTTT            | TACCAAGCCCA   | CCCAAT 300            |                 |
| Db | 1204217 | TTGATTAATTA   | TTCACGTA     | TCTTGAAAAAT              | CAGCCCTTTTCCGTCCGCTAGCA    |               | 1204158               |                 |
| QY | 301     | CAATGCCAGT    | ATATATG      | CCATTAAC                 | CGGCCCAAGCCCAT             | TAATCTGCAAT   | GTGTGA 360            |                 |
| Db | 1204157 | CAATGCAAT     | TTTACCTCA    | TATA-----AAGCCCAAGCCCAAT | CAGCAACTGTACT              |               | 1204107               |                 |
| QY | 361     | GCGATGTTG     | TGACCTTT     | TGATCA                   | TATAATATG                  | CCCTGTATCAAT  | TATCATATATA 420       |                 |
| Db | 1204106 | CCCCACAT      | TTTGCTTTAG   | AGTCAATTA                | TCCAAAGATGCCAT             | TTAGCTTGATGCA | CTMA 1204047          |                 |
| QY | 421     | TTCA          | CAGCGATG     | CCGTGAC                  | CTTTGA                     | AAAGTCTTAAGG  | ATCAAGCATG            | GCATCATGCG 480  |
| Db | 1204046 | TTGAA         | AAAGATGAT    | CTTAACCTTT               | GAAATAGCA                  | AGTCCGCTAGCA  | TATGATTTAATTT 1203987 |                 |
| QY | 481     | AAGATG        | CAGACCTG     | GCCTAAG                  | CGCAAGTGC                  | ATAGCATTTG    | ATAGCTTATG 540        |                 |
| Db | 1203986 | AATAC         | CTATAGCTTG   | AGCAATG                  | CTGTGTCG                   | CAAAATGTTCA   | TATATATATG            | CGCAATC 1203927 |
| QY | 541     | TTTGATTTT     | TATGATG      | ATGTTG                   | TATAAAGTTT                 | TTTTTGATGAT   | TATATATG              | CCATATCC 600    |

|    |                    |                        |            |
|----|--------------------|------------------------|------------|
| 1  | PRICE              | FLING DATE:            | 1995-04-21 |
| 2  | NUMBER OF SEQ      | ID NOS:                | 1          |
| 3  | SOFTWARE:          | PatentIn version       | 3.1        |
| 4  | SEQ ID NO 1        |                        |            |
| 5  | LENGTH:            | 1830121                |            |
| 6  | TYPE:              | DNA                    |            |
| 7  | ORGANISM:          | Haemophilus influenzae |            |
| 8  | FEATURE:           |                        |            |
| 9  | NAME/KEY:          | misc_feature           |            |
| 10 | LOCATION:          | (4747)..(4747)         |            |
| 11 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 12 | FEATURE:           |                        |            |
| 13 | NAME/KEY:          | misc_feature           |            |
| 14 | LOCATION:          | (9921)..(9921)         |            |
| 15 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 16 | FEATURE:           |                        |            |
| 17 | NAME/KEY:          | misc_feature           |            |
| 18 | LOCATION:          | (10150)..(10150)       |            |
| 19 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 20 | FEATURE:           |                        |            |
| 21 | NAME/KEY:          | misc_feature           |            |
| 22 | LOCATION:          | (29298)..(29298)       |            |
| 23 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 24 | FEATURE:           |                        |            |
| 25 | NAME/KEY:          | misc_feature           |            |
| 26 | LOCATION:          | (36543)..(36543)       |            |
| 27 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 28 | FEATURE:           |                        |            |
| 29 | NAME/KEY:          | misc_feature           |            |
| 30 | LOCATION:          | (36551)..(36551)       |            |
| 31 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 32 | FEATURE:           |                        |            |
| 33 | NAME/KEY:          | misc_feature           |            |
| 34 | LOCATION:          | (36636)..(36636)       |            |
| 35 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 36 | FEATURE:           |                        |            |
| 37 | NAME/KEY:          | misc_feature           |            |
| 38 | LOCATION:          | (40808)..(40810)       |            |
| 39 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 40 | FEATURE:           |                        |            |
| 41 | NAME/KEY:          | misc_feature           |            |
| 42 | LOCATION:          | (44416)..(44416)       |            |
| 43 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 44 | FEATURE:           |                        |            |
| 45 | NAME/KEY:          | misc_feature           |            |
| 46 | LOCATION:          | (44905)..(44905)       |            |
| 47 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 48 | FEATURE:           |                        |            |
| 49 | NAME/KEY:          | misc_feature           |            |
| 50 | LOCATION:          | (45593)..(45593)       |            |
| 51 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 52 | FEATURE:           |                        |            |
| 53 | NAME/KEY:          | misc_feature           |            |
| 54 | LOCATION:          | (45732)..(45732)       |            |
| 55 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 56 | FEATURE:           |                        |            |
| 57 | NAME/KEY:          | misc_feature           |            |
| 58 | LOCATION:          | (47036)..(47036)       |            |
| 59 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 60 | FEATURE:           |                        |            |
| 61 | NAME/KEY:          | misc_feature           |            |
| 62 | LOCATION:          | (51334)..(51334)       |            |
| 63 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 64 | FEATURE:           |                        |            |
| 65 | NAME/KEY:          | misc_feature           |            |
| 66 | LOCATION:          | (51602)..(51602)       |            |
| 67 | OTHER INFORMATION: | n equals a,t,c, or g   |            |
| 68 | FEATURE:           |                        |            |
| 69 | NAME/KEY:          | misc_feature           |            |

|   |               |
|---|---------------|
| LOCATION: (51786) .. (51786)              | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (51805) .. (51805)              | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (55369) .. (55369)              | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (65309) .. (65309)              | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (65313) .. (65313)              | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (80024) .. (80024)              | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (100091) .. (100091)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (102696) .. (102696)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (105121) .. (105121)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (107248) .. (107248)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (117136) .. (117136)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (119924) .. (119924)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (120038) .. (120038)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (121344) .. (121344)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (122167) .. (122167)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (122336) .. (122336)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (131340) .. (131340)            | a, t, c, or g |
| OTHER INFORMATION: n equals a, t, c, or g |               |
| FEATURE:                                  |               |
| NAME/KEY: misc_feature                    |               |
| LOCATION: (131360) .. (131360)            | a, t, c, or g |

```

1 OTHER INFORMATION: n equals a,t,c, or g
2 FEATURE:
3 NAME/KEY: misc_feature
4 LOCATION: (139910) ..(139910)
5 OTHER INFORMATION: n equals a,t,c, or g
6 FEATURE:
7 NAME/KEY: misc_feature
8 LOCATION: (140398) ..(140398)
9 OTHER INFORMATION: n equals a,t,c, or g
10 FEATURE:
11 NAME/KEY: misc_feature
12 LOCATION: (142750) ..(142750)
13 OTHER INFORMATION: n equals a,t,c, or g
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: (145058) ..(145058)
17 OTHER INFORMATION: n equals a,t,c, or g
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: (145171) ..(145171)
21 OTHER INFORMATION: n equals a,t,c, or g
22 FEATURE:
23 NAME/KEY: misc_feature
24 LOCATION: (145942) ..(145942)
25 OTHER INFORMATION: n equals a,t,c, or g
26 FEATURE:
27 NAME/KEY: misc_feature
28 LOCATION: (147197) ..(147197)
29 OTHER INFORMATION: n equals a,t,c, or g
30 FEATURE:
31 NAME/KEY: misc_feature
32 LOCATION: (150841) ..(150841)
33 OTHER INFORMATION: n equals a,t,c, or g
34 FEATURE:
35 NAME/KEY: misc_feature
36 LOCATION: (152500) ..(152500)
37 OTHER INFORMATION: n equals a,t,c, or g
38 FEATURE:

```

|                           | Score | DB                | Length            |
|---------------------------|-------|-------------------|-------------------|
| Query Match               | 11.6% | 160.4             | 1830121           |
| Best Local Similarly      | 49.6% | Pred. No. 1.4e-31 |                   |
| Matches 683; Conservative | 0     | Mismatches 606;   | Indels 89; Gaps 7 |

|    |         |  |         |
|----|---------|--|---------|
| Qy | 1       | TAAGCGTGAACCATGTTAACAACCTATACCGCTGCTATTATTAATTTTAAATTGATC      | 60      |
|    |         |  |         |
| Db | 1204439 | TAAACATTGAGCTAAATGCGTAAATTTCTTGCGCGGCTTTCAAAAGACAACTGATC       | 1204380 |
| Qy | 61      | CATCTTAGCACAAGCAGGTGACAAAGACAGCCCTTAGCGGTGCTTGACGTGCGCA        | 120     |
|    |         |  |         |
| Db | 1204379 | GAGCTTGACACAGCAGGCGACAAATATACCA-----TATCTCGCTTGCA              | 1204332 |
| Qy | 121     | TACTGCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAATGTGCATCCTTT       | 180     |
|    |         |  |         |
| Db | 1204331 | TGTTGGGGGTAAATAATCTATCGCTGTGCATGTATCGAACAA-----ATTACTTTG       | 1204278 |
| Qy | 181     | GAGTAAATCTTGCTCAATTTATCGGGGACCTTTGACCGATTTAAATACACTATTCACCCA   | 240     |
|    |         |  |         |
| Db | 1204277 | CGATGAATAATTTTGCAAGCAGCGCACCATCTCGACCAAAACAATATACAAATATATGTGTG | 1204218 |
| Qy | 241     | CTGATAAACATATGTCTGACAGCTCACCTAAATCTGTACCTTTACCAAGCCACCCAAAT    | 300     |
|    |         |  |         |
| Db | 1204217 | TTGATTAATTAATTCACCTAATCTGAAAAATCAGCCCTTTTCCGTCCGCTACCA         | 1204158 |
| Qy | 301     | CAATGCCAGTATGATATGCACTAAACCGTGCCAAAGCCATCAATTGCTGCAATGTGTGA    | 360     |
|    |         |  |         |
| Db | 1204157 | CAATTCGAATTTACCTTCATATB-----AAGCCCAAGCCAAATCAGCAACTGTACT       | 1204107 |
| Qy | 361     | GCCGATGTTTGTACCTTTTGATCATTAATAATATGCTTGATTCATATATCATCAATATA    | 420     |
|    |         |  |         |
| Db | 1204106 | CCCAACATTTTGCTTTAGTGATTAATCCAAAGAAATGCAATTAAGTATGATGACATTA     | 1204047 |
| Qy | 421     | TTCAACGATGCGGTAGCCCTTTGAAAGTCTTAAGGTATCAAGATGCAATGCATTCATGGG   | 480     |
|    |         |  |         |

|   |         |  |         |
|---|---------|--|---------|
| Dh  | 1204046 | TTGAAAGACGTATCTAACCCCTTTGAAATAGAGAAAGTGGCGATGACGAATTGAATCTAAATT  | 1203987 |
| Qy  | 481     | AAGATGGCACTGTGTCCTTAAGGCAAGTGCAGATAGAGCAATTGAGTGGTTATGCTTGGC     | 540     |
| Dh  | 1203986 | AATACCTATAGCTTGTGCCAAATGCTGTGTCTGCCAAATATGTTCAATATATTTAGCGACG    | 1203927 |
| Qy  | 541     | TTTGAATTTTAAATGATGGAATTGTGTAAGTTGTTTTTTTGATGATATATATGCCATACC     | 600     |
| Dh  | 1203926 | AACCAATGTAGCTTCTTCACAAGGTAAATATCACTTCATCTTTTACCATTTAAATATGTCTT   | 1203867 |
| Qy  | 601     | ATCACCATCATCAATATAAATAAAATCTGC--CGTTTGGTGGCTAAGTATGCTATTTGTTG    | 659     |
| Dh  | 1203866 | GCCATTTTTCAGTTTTTATGGCAATTAATCCGACATATTTTCCGAAAGAAACGGTATGTTT    | 1203807 |
| Qy  | 660     | TGATTAATGATGTGATTTTGGGGCAAGCGTTTGCAGTGTGTGTCAGCAATTCCTTGGCTAAAT  | 719     |
| Dh  | 1203806 | CGCTTGATTTTGTGTTTTTCCCAAAAGTCAGCC-----                           | 1203775 |
| Qy  | 720     | TGGCATCGTCTTGTGATGATATGCGGTTTGGCAATTTTCAAAATATAGCAATTTTGTG       | 779     |
| Dh  | 1203774 | ---TATCTTCATTTGTTCAACAACACTACTTTCAGCATATGATATTAAGCTAATTTTGCTT    | 1203718 |
| Qy  | 780     | CCAAATATATCCACCATGCCATCGTCCGATTCAGATGATATACCGAGAAATTTAAATTG      | 839     |
| Dh  | 1203717 | GCGCAATATCTTCTTAATATCATATATAGCAATATGATATGATCTTTCAGTCAAGTTCACACAG | 1203658 |
| Qy  | 840     | TTGCCCTTTGAGCTGCCAAATTTGAGATATATGCTCAAGCTGAAACCTTGCACAGTCAAGCA   | 899     |
| Dh  | 1203657 | TCCGTGCGGACCTTTTAAAGCTATATAGTGTCTCAAGCTGAAATCTAGAAAGCTCTAGTA     | 1203598 |
| Qy  | 900     | CCACCAATATCCATATTTTGGATGCTTAAATATTTCAAGTGCAGGCGTGCATATTTACCGC    | 959     |
| Dh  | 1203597 | CATAAGTTTCAAACTTTCATTC--AACAAATGACAAAGCGGAAATCCCAATATTTCCCG      | 1203541 |
| Qy  | 960     | CGACACCAACAAATATATGCTGATGTTTGGCCATCTGCGCTATTAATGTCTATACAGTGC     | 1019    |
| Dh  | 1203540 | CCATATCCAACTTTCAACACGACGCTTGTGCAATTTCAATAACTAAAGTATACGGTAC       | 1203481 |
| Qy  | 1020    | TTTTGGCATTTGAGCTGTGATGAGCGATATATGGTGATGTTTGTGCTTGAATTAATCTC      | 1079    |
| Dh  | 1203480 | TTTTTACATTTGACACTGTATATCCCAACAATTGGCTTGTCCCTGCGCGGCA-----        | 1203429 |
| Qy  | 1080    | GTCCTTGTAGTATCAATGAACAAATGGAACATGCTGTATCACAGAAATGCTTGAGCTT       | 1139    |
| Dh  | 1203428 | -----GAATTAATTAATATGCGCATTACTTCCACTCCGCGTTTAA                    | 1203388 |
| Qy  | 1140    | TGCGCGGCACAATATCTCGCGCTTCTTGGGTTAATGCGGGGCTGATGATGATCTGATCGG     | 1199    |
| Dh  | 1203387 | GTCGCGTTTGAATTTCTGTGTGTTTTTATCCGCAAGCCTGTGGCTAATTAACATCATATCCG   | 1203328 |
| Qy  | 1200    | CTGAGCTTAAGAGTTCACGTGTCCAAAGCTGCCAAATATGACAAATGAAACCGCTGTGGGAGTT | 1259    |
| Dh  | 1203327 | TTTCAAGTAAACCATTCCTGATTTTAACTATCACATGATGAAAGAGGATATTTTGAAGAAATT  | 1203268 |
| Qy  | 1260    | TATCTGCCAAGTAGATTTGGGATTAACATTCAGATGTGCTAATTTTGTGACCGTGAATTA     | 1319    |
| Dh  | 1203267 | TATCAATATCCAGTAGATTTTTTTTGCAGTATCAATACAGAAATATTAGCGCTGTGGGATA    | 1203208 |
| Qy  | 1320    | CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCAAACCGACGACGAGATATTTTTT        | 1377    |
| Dh  | 1203207 | AGAGATATATCCACACAGAAAGACCTGTTTTTCCAAAGCCCGATGATATGATATTTTTT      | 1203150 |
| RESULT 12   |         |  |         |
| US-10-795-159-569   |         |  |         |
| ; Sequence 569, Application US/10795159   |         |  |         |
| ; Publication No. US20050221439A1   |         |  |         |
| ; GENERAL INFORMATION:  |         |  |         |
| ; APPLICANT: BAKULETZ et al.  |         |  |         |
| ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE |         |  |         |
| ; FILE REFERENCE: 28335/38815A  |         |  |         |
| ; CURRENT APPLICATION NUMBER: US/10/795,159   |         |  |         |
| ; CURRENT FILING DATE: 2004-03-05   |         |  |         |

| Query Match  | Similarity  | Score           | DB | Length                            |
|--|---|-----------------|----|-----------------------------------|
| Best Local   | 49.0%   | Pred. No. 4e-31 |    |                                   |
| Matches  | 591   | Conservative    | 0  | Mismatches 346; Indels 69; Gaps 4 |
| PRIOR APPLICATION NUMBER: US 60/453,134<br>PRIOR FILING DATE: 2003-03-06<br>NUMBER OF SEQ ID NOS: 771<br>SOFTWARE: Patentin version 3.2<br>SEQ ID NO 569<br>LENGTH: 14324<br>TYPE: DNA<br>ORGANISM: H. influenzae<br>US-10-795-159-569 |   |                 |    |                                   |
| 172  | ATCACCTTTGAGTAATCTTGCCTCAATTATCGGGGCACTTTCAGACGATTAATAACACT   | 231             |    |                                   |
| 5280   | ATTAGCTTTGTGATGAAATTTTGCAGCTGCACACACTCTGCACCAAAACATTAACAAT    | 533             |    |                                   |
| 232  | ATCCACCACCTGCATACATATGCTGCACGCTCACTAAATCTTGACCTTTACCAAGCC     | 291             |    |                                   |
| 5340   | AATGTGTGTTGATTAATTAATTAATGAGCTAATTCTGAAAAATCACCCCTTTCCGTCYC   | 539             |    |                                   |
| 292  | ACCCAAATCAATGCCAGTGAATATGCCATAAACCGTGCACGCCCATCAATTGCTGC      | 351             |    |                                   |
| 5400   | GCCTAGCAACAAATGCAAT-----TTACCTCAATATPAAAGCCCATTAATGCAAGC      | 5450            |    |                                   |
| 352  | AATGTTGAGCGCAGTGTGTACCTTTTGAATCATTAATAATATGCTGTATCAATATC      | 411             |    |                                   |
| 5451   | AACGTACTCCCACTTTTGTGCTTTTGAATCATTAATCCAGAAATACATTAAGCTTG      | 5510            |    |                                   |
| 412  | ATCAATATATTCACAGCGATGCGGTAGCCCTTTGAAAGCTTAAAGGTATCAAGCATGC    | 471             |    |                                   |
| 5511   | ATGACATATATGAAAAAGATGATCTAACCTTTGAAATGACGAAGTGGGTACGATTTGA    | 5570            |    |                                   |
| 472  | ATCCATGGGAAGATTGGCAGCTGTGCTTAAGGCAAGCGCATTAAGCATTAAGCATTAAG   | 531             |    |                                   |
| 5571   | ATCTAATTAATTAACCGACAGCTGTGCTCAAGTGTGCTGCGCAAAATGTTTCATTAAT    | 5630            |    |                                   |
| 532  | ATGCTGCGCTTGAATTTTAAATGATGATTTGTGTAAGTTTGTGTTTGTGATATTA       | 591             |    |                                   |
| 5631   | ATGGCGACCGCATATGTACTTCTTCAAGGCAAAATTAATCTTATCTTTTACCATTA      | 5690            |    |                                   |
| 592  | TGCCATACCATCAACATCATCAATAATAATAAAATCTGCGTGTGTGCTAAGTAAGCT     | 651             |    |                                   |
| 5691   | GTATTTGCTGTCATTTTCAATTTTGTAGCAATATCCGACTATTTTCCGAAAAAAGAA     | 5750            |    |                                   |
| 652  | ATTGTTGTGATTAATGTTGATTTGGGGCAAGCTGTCAAGTGTGCTAAGCATGCTT       | 711             |    |                                   |
| 5751   | GGTTTGTCTCGCTGATTTTCACTTCAACCAAAAGTACG-----                   | 5790            |    |                                   |
| 712  | GGCTAATTTGGCATGCTGTGTGTAGATGATGGGGGTTTGGCAATTTTCAAAATACGGA    | 771             |    |                                   |
| 5791   | -----TTATCTCATATATCAAACTCCGACCTCAGCATTAATGATTAATGCGTAA        | 5840            |    |                                   |
| 772  | TTTTTGTGCCAAATATCAACCATGCGATGTCGCGATCAAGATGATCAGCGAGATTT      | 831             |    |                                   |
| 5841   | TTTTGCTTTGGGATTAATCTTAAATCAATTAAGCATCAATATGATCTTCAAGTCAAGT    | 5900            |    |                                   |
| 832  | TAAATATTTGGCCCTTGTAGCTGCACAAATTTGAGATATGCTCAAGCTGAAAACTTGAC   | 891             |    |                                   |
| 5901   | CAAAACAGTCCGACCGCGAGCTTTTAAAGCTAATTAAGTTGTCTCAAGCTGAAAACTTAGA | 5960            |    |                                   |
| 892  | CTCAAGCAACCAACCAATTCATTTTGGATGCTTAATTAATCAAGTCAAGCGCGCAAT     | 951             |    |                                   |
| 5961   | CTTAGTATCAATTAAGTTTCAATCTTCAATC--AACTATGACAAAGGGGAAATCCCAAT   | 6017            |    |                                   |
| 952  | ATTACCGCGCACCAACATCATGCTGCACTGTTTGGCATCTCGCTAATTAATGTCGT      | 1011            |    |                                   |
| 6018   | ATTTCGCGCATACCACTTTCAGCGACAGCGCTTGGCATTTCAATAAATTAAGTAT       | 6077            |    |                                   |
| 1012   | AACAGTGTGTTGGCATTTGAGCTGTGATGCGCATTAATTTGTGTGTTGCTTGAAGT      | 1071            |    |                                   |
| 6078   | AACAGTACTTTTTCATTTGAACTGTAAATCCCAACATTTGGCTTGTGCTGCGCG---     | 6134            |    |                                   |

|    |      |  |      |
|----|------|--|------|
| Oy | 1072 | TAGATCTGTCGCTTTGAGTGTATCAATGAACAATTGACATCGCTGATATACAGGAATGCC   | 1131 |
| Db | 6135 | -----GCAGAAATTAATTCAAATATGCGCATTAATCTTCACATCC                  | 6170 |
| Oy | 1132 | TTGAGCTTTGGCGGGGACATACTCGGCGCTTCTTGAGGTTAATGCCGGGGCTGATGATAT   | 1191 |
| Db | 6171 | CGCTTTAAGTCGGGTTTGATTTCTGAGTGTATTTTACCGCAGACCTTGGGCTTAATACAT   | 6230 |
| Oy | 1192 | CTGATCGGCTGAGCTTAAGAGTTCACTGTCCAACTGCCAAATGACAAATGAAGCGCTGT    | 1251 |
| Db | 6231 | CATATCGCTTCAAGTAACCATCTCTGATTTAACTACACAGTATGAAGGGATATTTGG      | 6290 |
| Oy | 1252 | GGGCGATTTATCTGCCAAGTNGAATTGGGATTAACCATGATGATGTAACTTTGTGACC     | 1311 |
| Db | 6291 | AGGAAGTTTATCAATATCACAGTAGGCTTTTTTGGGATCAATATCACAGAAATATTAGCCTG | 6350 |
| Oy | 1312 | GTCATTTGACCAAAAAATTACACGCTGACAGTCCAGAGCTGCCCAACCGACAGCGCATAT   | 1371 |
| Db | 6351 | TTGGGATTAAGATATATTCACACAAAGAAAGACCTGTTTTGCCAAGCCCAATATGCTAAT   | 6410 |
| Oy | 1372 | TTTTTT 1377  |      |
| Db | 6411 | ATTTT 6416   |      |

## RESULT 13

```

US-10-795-159-685/c
; Sequence 685. Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US-10/795,159
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 685

```

```

: LENGTH: 908766
: TYPE: DNA
: ORGANISM: H. influenzae
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9223)..(9223)
: OTHER INFORMATION: n is a, c, g, or t
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (39640)..(39640)
: OTHER INFORMATION: n is a, c, g, or t
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (292404)..(292404)
: OTHER INFORMATION: n is a, c, g, or t
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (908604)..(908641)
: OTHER INFORMATION: n is a, c, g, or t
US-10-795-159-685

```

|                           |        |                    |            |                |
|---------------------------|--------|--------------------|------------|----------------|
| Query Match               | 11.2%  | Score 154.4;       | DB 10;     | Length 908766; |
| Best Local Similarity     | 49.0%; | Pred. No. 4.4e-30; |            |                |
| Matches 591; Conservative | 0;     | Mismatches 546;    | Indels 69; | Gaps 4;        |

|    |        |   |        |
|----|--------|---|--------|
| Qy | 172    | ATCACCTTGAGTAAATCTTGCTCAATTATCGGGGACCTTGACCGATTAAATACACT  | 231    |
| Db | 107917 | ATAGCTTTGATGAAATTTTGCAAGCTGCACACCATCTCGACCAAAACATATCAAAAT | 107858 |

| QY | 232    | ATCCACCACGCAATACATATGCTGACAGCTCACTAAATCTTACCTTTCCAAAGCC      | 291     |
|----|--------|--|---------|
| QY | 107857 | AATGTGTGTTGATTAATTAATTAATCAGCTTAATCTGAAATATAGAGCCCTTTTCGCTCC | 1077988 |
| QY | 292    | ACCCAAATCAATGCGCAGTATGATATATGCAATAAACGTCGCAAGCCCATCAATGCTCG  | 351     |
| Db | 107797 | GCCTAGCAACAATGCAAT-----TTACCCCAATATATAAGCCGCAATATGACAG       | 107747  |
| QY | 352    | AATGTTAGCGGATGTTTGTACTTTTGAATCAATTAATAATAGCTTCTCAATATC       | 411     |
| Db | 107746 | AACGTACTCCCCACATTTGTGTCTTGAAGTCATTAAACCAAGAAATCCATTACCTTG    | 107687  |
| QY | 412    | ATCATATATTTACACAGCATGCGGTAGCCCTTTGAAGTCTTAAGGTAATCAAGCATGCG  | 471     |
| Db | 107686 | ATGCACTAATTTGAAAAACGTATCTTAACCTTTGAATGACGAAGTGCAGTACGAATTGA  | 107627  |
| QY | 472    | ATCCATGGGAAGATTGGCACTGTGCTTAAGGCAAGTGCAGATTAAGCAATTGAGTACGTT | 531     |
| Db | 107626 | ATCTAATTAATATCCGACAGCTGTGCCAAATCTGTCTGTGCCAAATATGTATATATAAT  | 107567  |

[illegible]

Oy 712 GGCCTAAATGGCATCGTCTTGCTAATGATGCGCGTTTGGCATTTTCAAAATAACGCAA 771  
 Db 107406 -----TTATCCTCATTTATCAAACTCCGACCTCAGCATTTAGTAATAATGCGTAA 107357  
 Oy 772 TTTTGTGCGCAATATCATCACCATCGCATCGGCGATCAAGTATGTACGCGAGAGATT 831  
 Db 107365 TTTTCTCTGCGCATTAATCTTTCTTAATTCATATATAGGCAATTCCTATATATCTTCTGCTGCGTT 107287

1072307 TAAATTTGTTGGCCCTTGAGCTGCCTAAATTTGAGTATGCTCAAGGTGAAAACTTGACG 891  
 QY TAAATTTGTTGGCCCTTGAGCTGCCTAAATTTGAGTATGCTCAAGGTGAAAACTTGACG 891  
 1072306 CAAAACAGTCGAGCGCGCAGCTTTTAAAGCTATGAAGTGTCTCAAGGTGAAAACTTAAAGG 107237  
 Db CAAAACAGTCGAGCGCGCAGCTTTTAAAGCTATGAAGTGTCTCAAGGTGAAAACTTAAAGG 107237  
 QY TCAAGCACCAACCAATCCATATTTTGGATCGTTATAATTCAGTGACGGCGTCGCAAT 951  
 892 TCAAGCACCAACCAATCCATATTTTGGATCGTTATAATTCAGTGACGGCGTCGCAAT 951

|    |        |   |        |
|----|--------|---|--------|
| Db | 107236 | CCTCTAGTCATCAAACTTCAGATCTTTCATC---AACATGACAAAGCCGGAAATCCCAAT  | 107180 |
| Qy | 952    | ATTACCGCGCAGACCAACAATCATGCTCCGTCATGTTTGGCATCTCGCCACTAATATGCTG | 1011   |
| Db | 107179 | ATTTCGGCCCATACCACTTTACCGCCAGCACTTTGGCATTTCATAAATTAATTAAGT     | 107130 |
| Qy | 1012   | AACACTGCTTTTGGCATTTGACCTGTGATGGCCAGATTTGGAGCTTTGGTCTTGAGT     | 1071   |

|    |        |  |        |
|----|--------|--|--------|
| Db | 107119 | TACGAGACTTTTTCATTGAACTGTATATCCCACTATGGCTTTGTGCTGCGG----    | 107063 |
| Qy | 1072   | TGATCTGTGCTTGAAGTGTTCATCAATGAACATTTGACATGCTGTATACAGAAATGCC | 1131   |
| Db | 107062 | -----GCGATTAATTCATAATATGCGCGATTAATCTTCCACTCC               | 107027 |
| Qy | 1132   | TTGACCTTTGGCGGCGACATACTCGGCGTTCTTGGGTTAATGCCGGGCTGTATGAT   | 1191   |

|    |        |  |        |
|----|--------|--|--------|
| Db | 107026 | CGCTTTAAGTGGGGTTTGATTTTCGGTGTTTTTACCGCAACCCCTGGGCTATAATCAAT    | 106967 |
| Qy | 1132   | CTGATCGGCTGAGCTTAAGATTCACCTGCAACGCTGCCAATAATGCAATGAACGGCTGT    | 1251   |
| Db | 106966 | CATATCGCTTCAAGTAACCATTTCCGATTTTAACTATACAGATATGAAGGAGTATTTTG    | 106907 |
| Qy | 1252   | GGGCACTTATTTGGCCAAAGCTAGGATTTGGGATTTACCATGATGATGATCTACTTTGGACC | 1311   |

Db 106906 AGGAGTTATCAATATCCAGTAGGCTTTTTCGGGTATCATTCACACGAAATATTAGCCTG 106847



Qy 1312 GTGATTGACCAAAATTCACAGCTGACAGCTCCAGAGCCGCAAAACGAGCATGATA 1371  
|||  
Db 106846 TTGGGATTAAGAGTAATATTCACACAGAAAGACCTGTTTCCCAAGCCCAATGATCGTAAT 106787  
Qy 1372 TTTT 1377  
|||  
Db 106786 ATTTT 106781

## RESULT 14

US-10-282-122A-25165/c  
; Sequence 25165, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITPA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25165  
; LENGTH: 1281  
; TYPE: DNA  
; ORGANISM: Legionella pneumophila  
US-10-282-122A-25165

Query Match 11.4%; Score 152.8; DB 8; Length 1281;  
Best Local Similarity 51.6%; Pred. No. 2.8e-31;  
Matches 464; Conservative 0; Mismatches 412; Indels 24; Gaps 4;

Qy 170 GCATCAGCTTGGTAATATCTTGCTCAATATGCGGGCACTTGACCGAATTAATACACA 229  
|||  
Db 1163 GGAACACTTGGTGAAGCGAATCTTGTGACGACCTTCACCAATCAAAACATG 1104  
|||  
Qy 230 CTATCCACCCAGCTGCAATACATATGCTGACAGCTCACTAAATCTTGACCTTAACGAGC 289  
|||  
Db 1103 GAACGAACAATCTGAAACTGGCTGGCTAATCTTGGAAATCCGCCCTTTTCCCTGT 1044  
|||  
Qy 290 CCAACCAAAATCAATGCGCACTGATGATGCAATAACCGTGCAGAGCCCATCAATGTCT 349  
|||

Db 1043 CCAACCCGAATAGAACAAATTTCCCTGCAATGAAC---TCCTATTCATTAATAGCA 987  
Qy 350 GCATAGGTGAGCGCAATGTTTGAACCTTTTGAATCAATTAATAATATGCTGTATCAATA 409  
|||  
Db 986 GAGATTGTAGCAACCAATATTTGTTCCCTTGAATCAATTAATCAACCTCATCTACT 927  
|||  
Qy 410 TCATCAATATATACAGCAGATGCGGTAGCCCTTTAAAGTCTTAAGGGATCAACATG 469  
|||  
Db 926 TCTCTTACCACCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867  
|||  
Qy 470 GCATCCATGGAAGATTTGGACGCTGTGCTTAAGGCAAGTGCAGATAAGGATTAAGTAG 529  
|||  
Db 866 TGTTCATAGAAATTCCTCGCTGCTGCTTAAGGCAAGGATTAAGGATTAAGTAG 807  
|||  
Qy 530 TTATGCTGCTTGTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 589  
|||  
Db 806 TTATGACCCCTTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 747  
|||  
Qy 590 AATGCAATACCATCATCATCATCAATTAATAATAATAATAATAATAATAATAATAATA 649  
|||  
Db 746 AATATAGTATATCTCTTCTTCAATCAGCCCAATTTCCAT-----GAAG 698  
|||  
Qy 650 CTATTTGTTGTATTAATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 709  
|||  
Db 697 GCGGCTCTTGGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642  
|||  
Qy 710 TTGCTTAATATGCAATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 769  
|||  
Db 641 ---ACAGTGTAGACGCTTTCACATTAATAATAATAATAATAATAATAATAATAATA 585  
|||  
Qy 770 AATTTTGTGCGCAATATATCAATCATGCAATGCTGCGCAATCAATGATCAAGCGAGAGA 829  
|||  
Db 584 TGTTTTCTGCTGCTATATGCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 525  
|||  
Qy 830 TTTAAATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889  
|||  
Db 524 TTAAGATGCTGCAACCACTGCGCAAAAGATTAAGTAAATCCAATTAAGTAAATCTGAT 465  
|||  
Qy 890 AGCTCAAGCACCACCAATC---CATTTTGTGATGATGATGATGATGATGATGATGATGAT 946  
|||  
Db 464 AACTCAATATCCCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405  
|||  
Qy 947 CCAATATTAACCGCGACACCAACATCATGCTGCAATGTTTTCATCTGCTACTAAT 1006  
|||  
Db 404 CCGATATTTCCGGCTACAGTACGCAAAACCGGCGCTTACCATTTTCCACCAAA 345  
|||  
Qy 1007 GTGTAACAGTCTTTTGGCAATTTGAGCCTGTGATGCGCATGATGATGATGATGAT 1066  
|||  
Db 344 GTGTTACAGTGGATTTTACATTAAGTTCCGTTATGCAATGACAGGGGGCTAATCTCT 285  
|||

## RESULT 15

US-10-958-216-402/c  
; Sequence 402, Application US/10958216  
; Publication No. US20050181388A1  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, ALBED  
; APPLICANT: DHARMASI, AKIL  
; APPLICANT: VEDADI, MASOUD  
; APPLICANT: ALAM, MUHAMMAD ZAHOR  
; APPLICANT: AROSMITH, CHERYL  
; APPLICANT: AMREY, DONALD E.  
; APPLICANT: BEATTIE, BRYAN  
; APPLICANT: BUZADZITA, KRISTINA  
; APPLICANT: CANADIEEN, VERONICA  
; APPLICANT: DOMAGALA, MEGAN  
; APPLICANT: HOUSTON, SIMON  
; APPLICANT: KANAGARAJAH, DHUSHY  
; APPLICANT: LI, QIN  
; APPLICANT: MANSOURY, KAMRAN  
; APPLICANT: MCDONALD, MERRY-LYNN  
; APPLICANT: NETHERY, KATHLEEN  
; APPLICANT: NG, IVY

```

APPLICANT: OUYANG, HUI
APPLICANT: RINDER, BENJAMIN
APPLICANT: RICHARDS, DAMN
APPLICANT: TAI, MATTHEW
APPLICANT: THALAKADA, ROSANNE
APPLICANT: VALLEE, FRANCOIS
APPLICANT: VIRAG, CRISTINA
TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM BACTERIA
FILE REFERENCE: IPT-205.01
CURRENT APPLICATION NUMBER: US/10/958,216
CURRENT FILING DATE: 2004-10-04
PRIOR APPLICATION NUMBER: PCT/CA03/00462
PRIOR FILING DATE: 2003-04-02
PRIOR APPLICATION NUMBER: 60/369,511
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 60/385,089
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/385,751
PRIOR FILING DATE: 2002-06-04
PRIOR APPLICATION NUMBER: 60/386,553
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,577
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,367
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,566
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/386,390
PRIOR FILING DATE: 2002-06-06
PRIOR APPLICATION NUMBER: 60/386,601
Remaindng Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1132
SOFTWARE: Patent Ver. 3.3
SEQ ID NO 402
LENGTH: 1314
TYPE: DNA
ORGANISM: Haemophilus influenzae
US-10-958-216-402

```

```

Query Match      11.0%; Score 152.4; DB 10; Length 1314;
Best Local Similarity 49.2%; Pred. No. 3,7e-31;
Matches 678; Conservative 0; Mismatches 611; Indels 89; Gaps 7;

QY 1 TAAAGCTGACCACTGTTAACTATGACCTGCTGCTTAAATTTTAAATGATC 60
DB 1308 TAAACATGAGCTTAATGCTTAATCTTGGCGGCTTTCAAAAGACAACTGATC 1249
QY 61 CATGCTAGCACAAGCAGGTGACAAAGACACAGCCCTAGCGGTGCTTGACTGCTGCAA 120
DB 1248 GAGACTTGACACAGGACGACCAATATACCA-----TATCTCGGCTTGCAA 1201
QY 121 TACTGCCGATCAAGGTACCACTGATATATTTTGTATGATGCAAAATGTGATACCTTT 180
DB 1200 TGTGGGCGTAATAATCTATGCTTGTTCATTTGTATGAAACA-----ATTAATTTTG 1147
QY 181 GAGTAATCTTGTCTCAATTTATGCGGGGATCTGACGATTAATAATACACATATCACCA 240
DB 1146 GATGAAATTTTGTGACAGCGGACCACTTCGACCAAAACAAATTAATATATGTGTG 1087
QY 241 CTGCATTAATATGCTGACAGCTCACTAAATCTTGAAGCTTTCACCAAGCCCAACCAAAAT 300
DB 1086 TTGATTATTAATTCAGCTATATCTGAAATAATCAAGCCCTTTTCGCTCCGCTAGCAA 1027
QY 301 CAATGCCAGTATGATATGCCATTAACCGTGCCAGCCCATCAATGCTGCTGATGTTGA 360
DB 1026 CAATGCAATTAATCCTCAATATA-----AAGCCCAAGCCCATCAGAACTGTACT 976
QY 361 GCGGATGTTGTACTCTTTGAAATCAATTAATAATGCTGCTATCAATATCATCAATATA 420
DB 975 CCCACATTTGTGCTTTTATGAGTCATTAATCAAGATGCAATTAAGCTTGTATGACTTAA 916
QY 421 TTGACAGCATGCGGTAGCCCTTTGAAAGTCTTAAGGATATCAAGCATGGCATCATGGG 480

```

```

DB 915 TTGAAAACGATGATCTAACCCCTTTGAAATGACGAATGGGATCGAATTAATCTAAAT 856
QY 481 AAGATTGGCAGCTGTGCTTAAGGCAAGTCAATTAAGCAATTAAGGATTAATGCTTGGC 540
DB 855 AATACCTATAGCTTGTGCAATGCTGTGTGCAAAATGTTCAATTAATTAATGGGACCC 796
QY 541 TTTGATTTTAAATGATGATGATGTAAGTTGTTTGTGATGATATATATGATATGCAATACC 600
DB 795 AACCAATGATGCTTCTTCAAGATTAATCACTTATCTTTTACATTAATTAATTAATGCTT 736
QY 601 ATGACCATCATCAATTAATTAATAATCTGCG-CTTTGTGCTGCTAAGTAACTAATTTGTTG 659
DB 735 GCCATTTTCAAGTTTATGACCAATTAATCCGACATATTTCCGCAAAAGAAACGGTATGTT 676
QY 660 TGTATATGATGATTTTGGGCAAGCTTGTGATGATGATGATGATGATGATGATGATGAT 719
DB 675 CGCTGATTTTCTTTTCCCAAAAGTCAAGCC----- 644
QY 720 TGGCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
DB 643 ---TATCTCATTTGTTCAACACACCTTATGATGATGATGATGATGATGATGATGATGAT 587
QY 780 CCAATTAATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839
DB 586 GGGCATATCTTCTTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 527
QY 840 TTGCCCCCTGAGTGGCAATTTGATATGCTCAAGCTGAAACCTTGACAGCTCAAGCA 899
DB 526 TCGCTCCGAGCTTTTAAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATG 467
QY 900 CCACCAATCATATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 959
DB 466 CATTAATGCTCAATCTCTATTC---AGCAATGACAAAGGGGAATCCCAATATTTCCGG 410
QY 960 CGACACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019
DB 409 CCAATACCACTTAATCTTAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATG 350
QY 1020 TTTTGGATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1079
DB 349 TTTTACCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 288
QY 1080 GTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
DB 297 -----GAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 257
QY 1140 TGGGCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
DB 256 GTGCGGTTGAAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 197
QY 1200 CTGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
DB 196 TTTCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 137
QY 1260 TATCTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
DB 136 TATCAATACAGAGAGATTTTTCGATGATGATGATGATGATGATGATGATGATGATGATG 77
QY 1320 CCAAAATTTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
DB 76 AGAGATATATCAACACAAGAAAGACCTGTTTTCGCAAGCCGATGATGATGATGATGATG 19

```

Search completed: May 18, 2006, 10:00:18  
Job time : 1704 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

## OM nucleic - nucleic search, using sw model

Run on: May 18, 2006, 09:34:46 ; Search time 51 Seconds  
(without alignments)  
1909.160 Million cell updates/sec

Title: US-10-672-787-35\_COPY\_11357\_12736  
Perfect score: 1380  
Sequence: 1 taagcgcgtgaccatgttaa.....acgacgcacatcttttga 1380

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 224314 seqs, 35277956 residues  
Total number of hits satisfying chosen parameters: 433684

Minimum DB seq length: 22  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Database :

Published Applications NA New:  
1: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
2: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                | Description        |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1          | 34.2  | 2.5         | 1302   | US-11-217-529-118    | Sequence 118, App  |
| 2          | 32.6  | 2.4         | 474    | US-11-217-529-2634   | Sequence 2634, App |
| 3          | 32.2  | 2.3         | 3411   | US-10-504-120-2      | Sequence 2, Appl1  |
| 4          | 31.8  | 2.3         | 516    | US-11-270-287-3      | Sequence 3, Appl1  |
| 5          | 31.8  | 2.3         | 1632   | US-11-217-529-77328  | Sequence 77328, A  |
| 6          | 31.6  | 2.3         | 789    | US-11-217-529-3777   | Sequence 3777, Ap  |
| 7          | 31.6  | 2.3         | 1125   | US-11-217-529-3776   | Sequence 3776, Ap  |
| 8          | 31.4  | 2.3         | 3653   | US-10-504-120-16     | Sequence 16, Appl  |
| 9          | 31.4  | 2.3         | 5026   | US-11-217-529-166180 | Sequence 166180,   |
| 10         | 31.2  | 2.3         | 1176   | US-11-217-529-734    | Sequence 734, App  |
| 11         | 31.2  | 2.3         | 1629   | US-11-217-529-1381   | Sequence 1381, App |
| 12         | 31.2  | 2.3         | 2385   | US-11-217-529-76422  | Sequence 76422, A  |
| 13         | 31.1  | 2.2         | 1077   | US-11-217-529-4571   | Sequence 4571, Ap  |
| 14         | 31.1  | 2.2         | 1593   | US-11-217-529-80459  | Sequence 80459, A  |
| 15         | 30.8  | 2.2         | 1722   | US-11-217-529-78136  | Sequence 78136, A  |
| 16         | 30.8  | 2.2         | 1722   | US-11-217-529-79486  | Sequence 79486, A  |
| 17         | 30.6  | 2.2         | 455    | US-11-217-529-3849   | Sequence 3849, Ap  |
| 18         | 30.4  | 2.2         | 459    | US-11-301-554-720    | Sequence 720, App  |
| 19         | 30.4  | 2.2         | 1353   | US-11-217-529-625    | Sequence 625, App  |
| 20         | 30.4  | 2.2         | 1353   | US-11-217-529-79251  | Sequence 79251, A  |
| 21         | 30.4  | 2.2         | 1896   | US-11-217-529-75758  | Sequence 75758, A  |
| 22         | 30.4  | 2.2         | 3540   | US-11-217-529-2192   | Sequence 2192, Ap  |
| 23         | 30.2  | 2.2         | 1692   | US-11-217-529-80706  | Sequence 80706, A  |
| 24         | 30.2  | 2.2         | 4899   | US-11-217-529-4664   | Sequence 4664, Ap  |
| 25         | 30.2  | 2.2         | 15456  | US-11-324-284-23     | Sequence 23, Appl  |

|   |    |      |     |        |   |                     |                    |
|---|----|------|-----|--------|---|---------------------|--------------------|
| C | 26 | 30   | 2.2 | 1665   | 7 | US-11-217-529-3737  | Sequence 3737, Ap  |
| C | 27 | 30   | 2.2 | 1764   | 7 | US-11-217-529-79615 | Sequence 79615, A  |
| C | 28 | 30   | 2.2 | 5778   | 7 | US-11-217-529-5615  | Sequence 5615, Ap  |
| C | 29 | 29.8 | 2.2 | 565    | 6 | US-10-488-619-1129  | Sequence 1129, Ap  |
| C | 30 | 29.8 | 2.2 | 567    | 6 | US-10-488-619-2704  | Sequence 2704, Ap  |
| C | 31 | 29.8 | 2.2 | 578    | 6 | US-10-488-619-1398  | Sequence 1398, Ap  |
| C | 32 | 29.8 | 2.2 | 582    | 6 | US-10-488-619-1435  | Sequence 1435, Ap  |
| C | 33 | 29.8 | 2.2 | 591    | 6 | US-10-488-619-1401  | Sequence 1401, Ap  |
| C | 34 | 29.8 | 2.2 | 656    | 6 | US-10-488-619-1128  | Sequence 1128, Ap  |
| C | 35 | 29.8 | 2.2 | 774    | 6 | US-10-488-619-1434  | Sequence 1434, Ap  |
| C | 36 | 29.6 | 2.1 | 1000   | 7 | US-11-233-726-4     | Sequence 4, Appl1  |
| C | 37 | 29.6 | 2.1 | 1182   | 6 | US-10-505-928-442   | Sequence 442, App  |
| C | 38 | 29.6 | 2.1 | 1182   | 6 | US-10-511-937-2869  | Sequence 2869, App |
| C | 39 | 29.6 | 2.1 | 1353   | 7 | US-11-217-529-2417  | Sequence 2417, Ap  |
| C | 40 | 29.6 | 2.1 | 2008   | 7 | US-11-246-999-26    | Sequence 26, Appl  |
| C | 41 | 29.6 | 2.1 | 2682   | 7 | US-11-217-529-432   | Sequence 432, App  |
| C | 42 | 29.6 | 2.1 | 3972   | 6 | US-10-473-173-131   | Sequence 131, App  |
| C | 43 | 29.4 | 2.1 | 2196   | 7 | US-11-217-529-1751  | Sequence 1751, Ap  |
| C | 44 | 29.4 | 2.1 | 4221   | 7 | US-11-217-529-2728  | Sequence 2728, Ap  |
| C | 45 | 29.4 | 2.1 | 128361 | 6 | US-10-505-928-151   | Sequence 151, App  |

## ALIGNMENTS

RESULT 1  
US-11-217-529-118/c  
; Sequence 118, Application US/11217529  
; Publication No. US20060099612A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNTORY LIMITED  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHIISA  
; APPLICANT: KODAMA, YUKIKO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIKO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: S-38-285  
; CURRENT APPLICATION NUMBER: US/11/217, 529  
; CURRENT FILING DATE: 2005-09-02  
; PRIOR APPLICATION NUMBER: US 10/932,182  
; PRIOR FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 118  
; LENGTH: 1302  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-11-217-529-118

Query Match  
Best Local Similarity 50.3%; Score 34.2; DB 7; Length 1302;  
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 446 AAGGTTAAGGATCAAGCATGATCATCGAAGATTGCGACCTGCTAAGCA 505  
DB 1147 AAGGCTAATGATTAAGAGAGGCTTAAGGCTTTAAATTGAGCTGCAAAAAGAC 1088

QY 506 AGTGCAGTAAGGATGAGTATGCTTGAATTTAATTGATGATGCT 565  
DB 1087 ATACCATGATTAACCAAGATCTGATCACTAGGAGGCTTCAATTGTAGATGGGT 1028

QY 566 AAGAGTTGTTTGTGATGATATATGCAATCAATCAGATCAATC 612  
DB 1027 AGATTGCTTAATAATATGAGGAATATGCTCAAAATTAATATTC 981

RESULT 2  
US-11-217-529-2634/c  
; Sequence 2634, Application US/11217529  
; Publication No. US20060099612A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNTORY LIMITED



```

RESULT 6
US-11-217-529-3777/c
Sequence 3777, Application US/11217529
Publication No. US20060099612A1
GENERAL INFORMATION:
APPLICANT: SUNTORY LIMITED
APPLICANT: NAKAO, YOSHIHIRO
APPLICANT: NAKAMURA, NORIHISA
APPLICANT: FUJIMURA, YUKITO
APPLICANT: KODAMA, YUKITO
APPLICANT: ASHIKARI, TOSHIHIRO
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
FILE REFERENCE: S-38-285
CURRENT APPLICATION NUMBER: US/11/217,529
CURRENT FILING DATE: 2005-09-02
PRIOR APPLICATION NUMBER: US 10/932,182
PRIOR FILING DATE: 2004-09-02
NUMBER OF SEQ ID NOS: 197023
SOFTWARE: PatentIn version 3.3
SEQ ID NO 3777
LENGTH: 789
TYPE: DNA
ORGANISM: Saccharomyces pastorianus
FEATURE:
NAME/KEY: modified base
LOCATION: (358)..(385)
OTHER INFORMATION: a, c, g, t, unknown, or other

```

|    | Query Match | Similarity   | 2.3%<br>49.4% | Score 11.6 | DB 7       | Length 1125 |
|----|-------------|--|---------------|------------|------------|-------------|
|    | Matches     | 82   | Conservative  | 0          | Mismatches | 84          |
|    |             |  |               |            | Indels     | 0           |
|    |             |  |               |            | Gaps       | 0           |
| Qy | 292         | ACCCAAATCATATCCGATGATGATATGCCATTAACCGTGCCAGGCCCATTAATGTGTC |               |            |            | 351         |
| Db | 369         | AGCCAATCCCTCCCAATGTTGTTCTTTCCAACTCATGAAGGAAATGATCTATC      |               |            |            | 310         |
| Qy | 352         | AATGTTGAGCCGATGTTTGACCTTTGTAATCATTTAAATATAGCCTTGCTATCAATTC |               |            |            | 411         |
| Db | 309         | AATGTTTTCATCTATGATGATCATCTGATCATCGCTGAATGCGCTCGAATGCTTTATC |               |            |            | 250         |
| Qy | 412         | ATCAATATATTCACAGGATGCGGTACCCCTTGAAGAGCTTAAG                |               |            |            | 457         |
| Db | 249         | ATTCAAGTTCAATAAAGATGATTCATCTCTGTAATACGCCATATAGG            |               |            |            | 204         |

RESULT 8  
US-10-504-120-16/c  
; Sequence 16, Application US/10504120  
; Publication No. US20060088829A1  
; GENERAL INFORMATION:  
; APPLICANT: Excellix, Inc.  
; TITLE OF INVENTION: M1NRs AS MODIFIERS OF INSULIN RECEPTOR SIGNALING AND METHODS OF  
; TITLE OF INVENTION: USE

FILE REFERENCE: EX03-003C-PC  
CURRENT APPLICATION NUMBER: US/10/504,120  
CURRENT FILING DATE: 2004-08-06  
PRIOR APPLICATION NUMBER: 60/354,824  
PRIOR FILING DATE: 2002-02-06  
PRIOR APPLICATION NUMBER: 60/358,217  
PRIOR FILING DATE: 2002-02-20  
PRIOR APPLICATION NUMBER: 60/358,189  
PRIOR FILING DATE: 2002-02-20  
PRIOR APPLICATION NUMBER: 60/358,126  
PRIOR FILING DATE: 2002-02-20  
PRIOR APPLICATION NUMBER: 60/358,995  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: 60/358,756  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: 60/358,765  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: 60/359,531  
PRIOR FILING DATE: 2002-02-25  
PRIOR APPLICATION NUMBER: 60/360,222  
PRIOR FILING DATE: 2002-02-26  
PRIOR APPLICATION NUMBER: 60/360,224  
PRIOR FILING DATE: 2002-02-26  
REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 16  
LENGTH: 3653  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-504-120-16

Query Match 2.3%; Score 31.4; DB 6; Length 3653;  
Best Local Similarity 56.2%; Pred. No. 5.2;  
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 753 AATTTCAAAATAGCGAATTTTGTGCAATATATCCACCATGCGATCAA 812  
Db 2893 AAGGTGTAAGATCCCAATTTGAAACCAATGCAACCCAGTCAAGTGA 2834  
Qy 813 GATGATCAGCGAGAGATTTAAATTTGTCCTTGAGCTGCCA 857  
Db 2833 GGATGATCAATGACTTTAAACCTTTGCTCCTTAAGTCA 2789

RESULT 9  
US-11-217-529-166180/c  
Sequence 166180, Application US/11217529  
Publication No. US20060099612A1  
GENERAL INFORMATION:  
APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA  
APPLICANT: KODAMA, YUKIKO  
APPLICANT: FUJIMURA, TOMOKO  
APPLICANT: ASHIKARI, TOSHIHIKO  
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
FILE REFERENCE: S-38-285  
CURRENT APPLICATION NUMBER: US/11/217,529  
CURRENT FILING DATE: 2005-09-02  
PRIOR APPLICATION NUMBER: US 10/932,182  
PRIOR FILING DATE: 2004-09-02  
NUMBER OF SEQ ID NOS: 197023  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 166180  
LENGTH: 5026  
TYPE: DNA  
ORGANISM: Saccharomyces pastorianus  
US-11-217-529-166180

Query Match 2.3%; Score 31.4; DB 7; Length 5026;  
Best Local Similarity 52.7%; Pred. No. 6.2;  
Matches 68; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 530 TTATGCTGCTTGTGATTTTAAATGATGATGTAAGATTGTTTTTGATATAT 589  
Db 3680 TTTTATTAGATAAATAATTTATTTACTCAGACATTAATAGTATTTAAATCTAT 3621  
Qy 590 AATGCCATACCATCACCATCATCAAAATTAATAAAAATCTGCCCTTGCTGAGTAAG 649  
Db 3620 AATTAATCATGATAAAGATCATGAAATTAATACAGATAATCCTATTATGCTTATATGT 3561  
Qy 650 CTATTTGTT 658  
Db 3560 TGGTTATTT 3552

RESULT 10  
US-11-217-529-734/c  
Sequence 734, Application US/11217529  
Publication No. US20060099612A1  
GENERAL INFORMATION:  
APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA  
APPLICANT: KODAMA, YUKIKO  
APPLICANT: FUJIMURA, TOMOKO  
APPLICANT: ASHIKARI, TOSHIHIKO  
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
FILE REFERENCE: S-38-285  
CURRENT APPLICATION NUMBER: US/11/217,529  
CURRENT FILING DATE: 2005-09-02  
PRIOR APPLICATION NUMBER: US 10/932,182  
PRIOR FILING DATE: 2004-09-02  
NUMBER OF SEQ ID NOS: 197023  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 734  
LENGTH: 1176  
TYPE: DNA  
ORGANISM: Saccharomyces pastorianus  
US-11-217-529-734

Query Match 2.3%; Score 31.2; DB 7; Length 1176;  
Best Local Similarity 50.0%; Pred. No. 3.2;  
Matches 78; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 429 GATCGGTAGCCCTTGAAGATCTTAAGGATATCAAGATGCGATCATGGAAGATTGG 488  
Db 634 GATCGATTTCTCTTGTAATTTCTCCAGATGTTAGCTTGAGCGGATTAATACCTGG 575  
Qy 489 CAGCTGCTTAAGCAAGTGCAGATPAAGCATGAGTTATGCTTCTTATTT 548  
Db 574 TAAATTTTGTTCAGTTCTCAGCTAATTCATCGACTGCTTCATCATATCTTCAAGTC 515  
Qy 549 TTAATGATGATGTAAGATTTGTTTTTGGAT 584  
Db 514 TTTTCTTTTCATTTCCAAAAGTTCTTTCTTGGCT 479

RESULT 11  
US-11-217-529-1381/c  
Sequence 1381, Application US/11217529  
Publication No. US20060099612A1  
GENERAL INFORMATION:  
APPLICANT: SUNTORY LIMITED  
APPLICANT: NAKAO, YOSHIHIRO  
APPLICANT: NAKAMURA, NORIHISA  
APPLICANT: KODAMA, YUKIKO  
APPLICANT: FUJIMURA, TOMOKO  
APPLICANT: ASHIKARI, TOSHIHIKO  
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
FILE REFERENCE: S-38-285  
CURRENT APPLICATION NUMBER: US/11/217,529  
CURRENT FILING DATE: 2005-09-02  
PRIOR APPLICATION NUMBER: US 10/932,182  
PRIOR FILING DATE: 2004-09-02

RESULT 15  
US-11-217-529-78136/c  
; Sequence 78136, Application US/11217529



```
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHITIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-265
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78136
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78136
```

```
Query Match      2.28; Score 30.8; DB 7; Length 1125;
Best Local Similarity 47.0%; Pred. No. 4.2;
Matches 95; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 518 GCATGAGTAGGTTAGTCCCTTGATTTTAAATTGATGATGATAAGTTGTT 577
Db 544 GCATTTTATAGCATGCGCTTCGTGACATTTTGAATTTTTCGTACTGTCAT 485
QY 578 TTTTGATGATATATGCGATACCATCACATCATCAATAATAAATCTGCCGTTGG 637
Db 484 TTCCTCTTACATATATCAAGTTTCAACCAATAACTTCATAAATCATTCAGAGACG 425
QY 638 TGGCTAAGTAGCTATTGTTGATATAGTGTGATTTGGGCAACGCTTGCAGTG 697
Db 424 TGCTGAAGTCATTTAAAGTAGATATGCTATTCTTATCTTGGTATGTTCAICTA 365
QY 698 GTCAGCATGCTTGCTAAT 719
Db 364 ATAAAAAAATCGTTCATTAAT 343
```

Search completed: May 18, 2006, 11:42:15  
Job time : 52 secs